

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:08:17 ; Search time 2.8 Seconds
(without alignments)
721.436 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRMHGLNHAFDQLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	351	2 A56387	helix-loop-helix t
2	84	76.4	312	2 A40708	basic-helix-loop-h
3	81	73.6	337	2 I48682	NEX-1 - mouse
4	81	73.6	337	2 I57038	gene Dlx-3 protein
5	81	73.6	352	2 I51687	neurogenic differe
6	80	72.7	347	2 JC7999	basic helix-loop-h
7	80	72.7	357	2 JC4703	basic helix-loop-h
8	80	72.7	357	2 I49338	neurogenic differe
9	80	72.7	381	2 A57059	beta-cell E-box tr
10	78	70.9	192	2 T15764	hypothetical prote
11	77	70.0	367	2 JC6087	helix-loop-helix t
12	73	66.4	71	2 T29378	LIN-32 protein - C
13	73	66.4	381	2 G02668	neurogenic basic-h
14	73	66.4	381	2 JC4647	KW8 protein - rat
15	73	66.4	383	2 JC4688	neuro D-related fa
16	71	64.5	237	2 T15548	hypothetical prote
17	70	63.6	326	2 S71755	transcription fact
18	67	60.9	207	2 T20391	hypothetical prote
19	65	59.1	334	2 G02409	protein kinase C-b
20	63	57.3	258	2 T21959	hypothetical prote
21	62	56.4	170	2 T29971	hypothetical prote
22	62	56.4	360	2 A56086	basic helix-loop-h
23	59	53.6	166	2 A33637	Xtvi protein - Afr
24	59	53.6	255	2 JN0624	Myogenic factor -
25	57	51.8	198	2 A57717	transcription fact
26	56	50.9	201	2 S23324	gene achaete prote
27	56	50.9	201	2 A43731	Achaete-scute comp
28	56	50.9	201	2 G01204	twist protein homo
29	56	50.9	206	2 I53066	gene M-twist prote

30 56 50.9 240 2 B43909 myogenic protein M
31 56 50.9 242 2 A34872 muscle-specific re
32 56 50.9 242 2 A42668 myogenic factor MR
33 56 50.9 242 2 S12385 myf-6 protein - hu
34 56 50.9 242 2 JC1233 muscle-specific re
35 56 50.9 260 2 A41123 myogenic transcrip
36 55 50.0 64 2 A48411 Myf5 homolog - chi
37 55 50.0 190 2 B43783 myogenesis protein
38 55 50.0 246 2 B43912 myogenic regulator
39 55 50.0 255 2 S16151 myogenic factor Xn
40 55 50.0 255 2 S22825 myf-5 protein - mc
41 55 50.0 255 2 S04105 myogenic factor 5
42 55 50.0 258 2 S41126 Myf5 protein - chi
43 55 50.0 289 2 A34783 myogenesis protein
44 55 50.0 297 2 A35874 myogenic factor 1
45 55 50.0 298 2 A32872 myogenic factor CM

ALIGNMENTS

RESULT 1

A56387

helix-loop-helix transcription factor MATH-1 - mouse

N;Alternate names: atonal homolog MATH-1

C;Species: Mus musculus (house mouse)

C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999

C;Accession: A56387

R;Akazawa, C.; Ishibashi, M.; Shimizu, C.; Nakanishi, S.; Kageyama, R.

J. Biol. Chem. 270, 8730-8738, 1995

A;Title: A mammalian helix-loop-helix factor structurally related to the product of Dros

A;Reference number: A56387; MUID:95238366; PMID:7721778

A;Accession: A56387

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-351 <AX>

A;Cross-references: GB:D43694; NID:g994770; PIDN:BAA07791.1; PID:d1008377; PID:g994771

C;Genetics:

A;Introns: #status absent

C;Keywords: transcription factor

Query Match 100.0%; Score 110; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 9.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21

|||||

Db 160 AANARERRMHGLNHAFDQLR 180

RESULT 2

A40708

basic-helix-loop-helix protein ato - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C;Accession: A40708

R;Jarman, A.P.; Grau, Y.; Jan, L.Y.; Jan, Y.N.

Cell 73, 1307-1321, 1993

A;Title: atonal is a proneural gene that directs chordotonal organ formation in the Dros

A;Reference number: A40708; MUID:93313961; PMID:8324823

A;Accession: A40708

A;Status: preliminary

A;Molecule type: nucleic acid

A;Residues: 1-312 <JAR>

A;Cross-references: GB:I36646; NID:9551565; PIDN:AAA21879.1; PID:9551566

A;Note: sequence extracted from NCBI backbone (NCBIN:135094, NCBI:135095)

C;Genetics:

A;Gene: FlyBase:ato

A;Cross-references: FlyBase:FBgn0010433

Query Match

Best Local Similarity 76.4%; Score 84; DB 2; Length 312;

Matches 81.0%; Pred. No. 1.1e-05;

```

Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1  AANARERRRMHGLNHFADQLR 21
      |||||  |||||  |||||  |||||
Db       259  AANARERRRMQNQAFLRLR 279

RESULT 3
148682
NEX-1 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Jun-2000
C:Accession: I48682; S69346
R:Bartholoma, A.; Nave, K.A.
Mech. Dev. 48, 217-228, 1994
A:Title: NEX-1: a novel brain-specific helix-loop-helix protein with autoregulation and
A:Reference number: I48682; MUID:95200803; PMID:7545978
A:Accession: I48682
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-337 <RES>
A:Cross-references: EMBL:U90986; NID:g881963; PIDN:AA014576.1; PID:g881962
R:Shimizu, C.; Akazawa, C.; Nakanishi, S.; Kagiyama, R.
Eur. J. Biochem. 229, 239-248, 1995
A:Title: MATH-2, a mammalian helix-loop-helix factor structurally related to the product
A:Reference number: S69346; MUID:95262673; PMID:7744035
A:Accession: S69346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <SHI>
A:Cross-references: EMBL:D44480; NID:g994772; PIDN:BAA07923.1; PID:g994773
C:Genetics:
A:Gene: nex-1

Query Match      73.6%; Score 81; DB 2; Length 337;
Best Local Similarity 80.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2  ANARERRRMHGLNHFADQLR 21
      |||||  |||||  |||||  |||||
Db       99  ANARERRRMHGLNDAFLNR 118

RESULT 4
157038
gene Dlx-3 protein - mouse
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C:Accession: I57038
R:Robinson, G.W.; Mahon, K.A.
Mech. Dev. 48, 199-215, 1994
A:Title: Differential and overlapping expression domains of Dlx-2 and Dlx-3 suggest dist
A:Reference number: I57038; MUID:95200802; PMID:7893603
A:Accession: I57038
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-337 <RES>
A:Cross-references: GB:S75299; NID:g896199; PID:g896200
C:Genetics:
A:Gene: Dlx-3

Query Match      73.6%; Score 81; DB 2; Length 337;
Best Local Similarity 80.0%; Pred. No. 3.5e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2  ANARERRRMHGLNHFADQLR 21
      |||||  |||||  |||||  |||||
Db       99  ANARERRRMHGLNDAFLNR 118

RESULT 5
151687
neurogenic differentiation factor - African clawed frog

```

F:102-157/Region: helix-loop-helix #status predicted

Query Match 72.7%; Score 80; DB 2; Length 357;
Best Local Similarity 80.0%; Pred. No. 5.4e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 106 ANARERNMHGHLNHAFFDQLR 125

RESULT 8

149338 neurogenic differentiation factor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I49338; I49345

R:Lee, J.E.; Hollenberg, S.M.; Snider, L.; Turner, D.L.; Lipnick, N.; Weintraub, H.

Science 268, 836-844, 1995

A:Title: Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix

A:Reference number: A56481; MUID:95273957; PMID:7754368

A:Accession: I49338

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 73-357 <RE2>

A:Cross-references: EMBL:U28068; NID:G854738; PIDN:AA52203.1; PID:G854739

A:Accession: I49345

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-121 <RE3>

A:Cross-references: EMBL:U28888; NID:G854742; PIDN:AA52204.1; PID:G854743

C:Genetics:

A:Gene: neuroD

Query Match 72.7%; Score 80; DB 2; Length 357;

Best Local Similarity 80.0%; Pred. No. 5.4e-05;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 106 ANARERNMHGHLNHAFFDQLR 125

RESULT 9

A57059

beta-cell E-box transcription activator 2 - hamster (fragment)

C:Species: Cricetinae gen. sp. (hamster)

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 08-Oct-1999

C:Accession: A57059

R:Naya, F.J.; Stellrecht, C.M.M.; Tsai, M.J.

Genes Dev. 9, 1009-1019, 1995

A:Title: Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix

A:Reference number: A57059; MUID:95293222; PMID:7774807

A:Accession: A57059

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-381 <NAY>

A:Cross-references: GB:U24679; NID:G777418; PIDN:AAA86518.1; PID:G777419

Query Match 72.7%; Score 80; DB 2; Length 381;

Best Local Similarity 80.0%; Pred. No. 5.8e-05;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 131 ANARERNMHGHLNHAFFDQLR 150

RESULT 10

T15764

hypothetical protein C34E10.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15764

R:Kirsten, J.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid C34E10.

A:Reference number: Z18399

A:Accession: T15764

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-192 <KIP>

A:Cross-references: EMBL:U10402; NID:G500723; PID:G500730; PIDN:AAA19069.1; CESP:C34E10.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C34E10.7

A:Introns: 65/2; 107/3

Query Match 70.9%; Score 78; DB 2; Length 192;

Best Local Similarity 75.0%; Pred. No. 5.8e-05;

Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 24 ANGRERARMHGLNVALDMLR 43

RESULT 11

JC6087

helix-loop-helix transcription factor, BETA3 - hamster

C:Species: Cricetinae gen. sp. (hamster)

C>Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999

C:Accession: JC6087

R:Pepton, M.; Stellrecht, C.M.M.; Naya, F.J.; Huang, H.P.; Samora, P.J.; Tsai, M.J.

Mol. Cell. Biol. 16, 626-633, 1996

A:Title: BETA3, a novel helix-loop-helix protein, can act as a negative regulator of B2

A:Reference number: JC6087; MUID:96140430; PMID:8552091

A:Accession: JC6087

A:Molecule type: mRNA

A:Residues: 1-367 <PEY>

A:Cross-references: GB:S80870; NID:G1911496; PIDN:AAB50691.1; PID:G1911497

A:Experimental source: insulin tumor cell

C:Comment: This factor belongs to the tissue-specific class B basic helix-loop-helix fam

ssion.

C:Genetics:

A:Gene: beta3

C:Keywords: repressor; transcription factor

Query Match 70.0%; Score 77; DB 2; Length 367;

Best Local Similarity 78.9%; Pred. No. 0.00016;

Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 NARERRRMHGLNHAFFDQLR 21
||||| ||||| ||||| |||||

DB 234 NARERRRMHGLNHAFFDQLR 252

RESULT 12

T29378

lin-32 protein - Caenorhabditis elegans

C:Alternate names: lin transcription factor

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T29378; S52415

R:Miller, N.; Langston, Y.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid T14F9.

A:Reference number: Z20615

A:Accession: T29378

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-71 <MIL>

A:Cross-references: EMBL:U50199; PIDN:AAA91264.1; CESP:lin-32

A:Experimental source: strain Bristol N2

R:Zhao, C.; Emmons, S.W.

Nature 373, 74-78, 1995

A:Title: A transcription factor controlling development of peripheral sense organs in C.
A:Reference number: S52415; MUID:95098129; PMID:7800042
A:Accession: S52415
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <ZHA>
A:Cross-references: EMBL:U15418; NID:G560021; PIDN:AAA67360.1; PID:G560022
C:Genetics:
A:Gene: CESP:lin-32
A:Introns: 23/3

Query Match 66.4%; Score 73; DB 2; Length 71;
Best Local Similarity 71.4%; Pred. No. 0.00013;
Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AANARERRRMEGLNHAFDQLR 21
|||||:|||||:
Db 5 AANERERRRMTNLNVAYDEL 25
|||||:|||||:

RESULT 13
G02668
neurogenic basic-helix-loop-helix (bHLH) protein - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C:Accession: G02668
R:Tapscott, S.J.; Tamimi, R.T.; McCormick, B.M.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01573
A:Accession: G02668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-381 <TAP>
A:Cross-references: EMBL:U58681; NID:G1477748; PID:G1477749
C:Genetics:
A:Gene: NeuroD2

Query Match 66.4%; Score 73; DB 2; Length 381;
Best Local Similarity 75.0%; Pred. No. 0.00074;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMEGLNHAFDQLR 21
|||||:|||||:
Db 126 ANARERNRMDLNAALDNL 145
|||||:|||||:

RESULT 14
JC4647
KW8 protein - rat
N:Alternate names: basic helix-loop-helix protein homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C:Accession: JC4647
R:Kume, H.; Maruyama, K.; Tomita, T.; Iwatsubo, T.; Saïdo, T.C.; Obata, K.
Biochem. Biophys. Res. Commun. 219, 526-530, 1996
A:Title: Molecular cloning of a novel basic helix-loop-helix protein from the rat brain.
A:Reference number: JC4647; MUID:96193685; PMID:8605021
A:Accession: JC4647
A:Molecule type: DNA
A:Residues: 1-381 <KUM>
A:Cross-references: DDBJ:D82868; NID:G1166397; PIDN:BAAL1615.1; PID:G1166398
A:Experimental source: brain
C:Comment: This protein is involved in synaptic plasticity, and has a role specific to a
-loop-helix domain.
C:Keywords: brain
F:122-134/Region: basic

Query Match 66.4%; Score 73; DB 2; Length 381;
Best Local Similarity 75.0%; Pred. No. 0.00074;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMEGLNHAFDQLR 21
|||||:|||||:
Db 126 ANARERNRMDLNAALDNL 145
|||||:|||||:

Db 126 ANARERNRMDLNAALDNL 145
|||||:|||||:

RESULT 15
JC4688

neuro D-related factor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C:Accession: JC4688
R:Yasunami, M.; Suzuki, K.; Maruyama, H.; Kawakami, H.; Nagai, Y.; Hagiwara, M.; Ohkubo,
Biochem. Biophys. Res. Commun. 220, 754-758, 1996
A:Title: Molecular cloning and characterization of a cDNA encoding a novel basic helix-
A:Reference number: JC4688; MUID:96183021; PMID:8607837
A:Accession: JC4688
A:Molecule type: mRNA
A:Residues: 1-383 <YAS>
A:Cross-references: DDBJ:D83507; NID:G1304163; PIDN:BAA11931.1; PID:G1304164
A:Experimental source: embryo
C:Comment: This factor plays distinct roles in neural development and plasticity as a

F:36-43/Region: proline-rich
F:83-92/Region: glutamic acid-rich
F:213-297/Region: alanine/glycine-rich

Query Match 66.4%; Score 73; DB 2; Length 383;
Best Local Similarity 75.0%; Pred. No. 0.00074;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ANARERRRMEGLNHAFDQLR 21
|||||:|||||:
Db 127 ANARERNRMDLNAALDNL 146
|||||:|||||:

Search completed: September 21, 2004, 21:25:00
Job time : 2.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:35:26 ; Search time 1.624 Seconds
(without alignments)
673.321 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRRMEGLNHFQDLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	351	1	ATH1_MOUSE
2	110	100.0	354	1	ATH1_HUMAN
3	84	76.4	198	1	AMOS_DROME
4	84	76.4	312	1	ATO_DROME
5	81	73.6	337	1	NDF6_MOUSE
6	81	73.6	337	1	NDF6_MOUSE
7	81	73.6	352	1	NDF1_XENLA
8	80	72.7	330	1	NDF4_MOUSE
9	80	72.7	330	1	NDF4_CHICK
10	80	72.7	331	1	NDF4_HUMAN
11	80	72.7	355	1	NDF1_MESAU
12	80	72.7	356	1	NDF1_HUMAN
13	80	72.7	357	1	NDF1_CHICK
14	80	72.7	357	1	NDF1_MOUSE
15	80	72.7	357	1	NDF1_RAT
16	78	70.9	192	1	YLB7_CAEEL
17	77	70.0	367	1	BET3_MESAU
18	76	69.1	315	1	NDF4_XENLA
19	73	66.4	71	1	LI32_CAEEL
20	73	66.4	382	1	NDF2_HUMAN
21	73	66.4	382	1	NDF2_RAT
22	73	66.4	383	1	NDF2_MOUSE
23	68	61.8	208	1	NGN1_BRARE
24	68	61.8	214	1	NGN3_MOUSE
25	67	60.9	214	1	NGN3_HUMAN
26	67	60.9	237	1	NGN1_MOUSE
27	67	60.9	244	1	NGN1_MOUSE
28	67	60.9	244	1	NGN1_RAT
29	67	60.9	250	1	NGN2_HUMAN
30	67	60.9	263	1	NGN2_MOUSE
31	66	60.0	398	1	TAP_DROME
32	65	59.1	298	1	OLG2_CHICK
33	65	59.1	323	1	OLG2_HUMAN

34	65	59.1	323	1	OLG2_MOUSE	Q9eqw6 mus musculus
35	62	56.4	360	1	DEI_DROME	P41894 drosophila
36	60	54.5	179	1	TF21_HUMAN	O43680 homo sapien
37	60	54.5	179	1	TF21_MOUSE	O35437 mus musculus
38	60	54.5	201	1	MUSC_MOUSE	O88940 mus musculus
39	60	54.5	206	1	MUSC_HUMAN	O60682 homo sapien
40	59	53.6	166	1	TWST_XENLA	P13903 xenopus lae
41	59	53.6	196	1	TWST_BRABE	O96642 branchiosto
42	59	53.6	255	1	MYP5_BOVIN	P17667 bos taurus
43	58	52.7	260	1	OLG1_MOUSE	Q9jkn5 mus musculus
44	58	52.7	261	1	OLG1_RAT	Q9wuq3 rattus norv
45	58	52.7	271	1	OLG1_HUMAN	Q8tak6 homo sapien

ALIGNMENTS

RESULT 1

ATH1_MOUSE						
ID	ATH1_MOUSE	STANDARD;	PRT;	351 AA.		
AC	P48985;					
DT	01-FEB-1996 (Rel. 33, Created)					
DT	01-FEB-1996 (Rel. 33, Last sequence update)					
DT	15-MAR-2004 (Rel. 43, Last annotation update)					
DE	Atonal protein homolog 1 (Helix-loop-helix protein MATH-1) (MATH1).					
GN	ATOH1 OR ATH1					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=129/J;					
RX	MEDLINE=95238366; PubMed=7721778;					
RA	Akazawa C., Ishibashi M., Shimizu C., Nakanishi S., Kageyama R.;					
RT	"A mammalian helix-loop-helix factor structurally related to the					
RT	product of Drosophila proneural gene atonal is a positive					
RT	transcriptional regulator expressed in the developing nervous					
RT	system.";					
RL	J. Biol. Chem. 270:8730-8738(1995).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=FVB/N; TISSUE=Colon;					
EX	MEDLINE=22388257; PubMed=12477932;					
RA	Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,					
RA	Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,					
RA	Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Haieh F.,					
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,					
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,					
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,					
RA	Raha S.S., Joquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,					
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,					
RA	Fabey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,					
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,					
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,					
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,					
RA	Butterfield J.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,					
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;					
RT	"Generation and initial analysis of more than 15,000 full-length					
RT	human and mouse cDNA sequences.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16999-16993(2002).					
CC	-!- FUNCTION: Activates E box-dependent transcription in collaboration					
CC	with E47, but the activity is completely antagonized by the					
CC	negative regulator of neurogenesis HES1. May play a role in the					
CC	differentiation of subsets of neural cells by activating E box-					
CC	dependent transcription.					
CC	-!- SUBUNIT: Efficient DNA binding requires dimerization with another					
CC	bHLH protein.					
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).					
CC	-!- TISSUE SPECIFICITY: Developing nervous system, and in adult					

epithelial cells of the gastrointestinal tract.

-!- DEVELOPMENTAL STAGE: First detected in the cranial ganglions and the dorsal part of the central nervous system on embryonic day 9.5 (E9.5). From E10.5 onward, prominent expression of MATH-1 continues in the dorsal part of the central nervous system but becomes restricted to the external granular layer of the cerebellum by e18 and is undetectable in the adult nervous system.

-!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; D43694; BAA07791.1; -;
 EMBL; BC010820; AAH10820.1; -;
 EMBL; BC051256; AAH51256.1; -;
 PIR; A56387; A56387.
 TRANSFAC; T01668; -;
 MGI; 104654; Atoh1.
 GO; GO:0007420; P:brain development; IMP.
 InterPro; IPR001092; HLH_basic.
 Pfam; PF00010; HLH; 1.
 SMART; SM00353; HLH; 1.
 PROSITE; PS00888; HLH; 1.
 Transcription regulation; Activator; DNA-binding; Nuclear protein.
 DNA BIND 157 168 BASIC DOMAIN
 FT DOMAIN 169 209 HELIX-LOOP-HELIX MOTIF.
 FT DOMAIN 221 224 POLY-PRO
 SQ SEQUENCE 351 AA; 37854 MW; 596E03DF2C3BF96 CRC64;

Query Match 100.0%; Score 110; DB 1; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
 |||||
 DB 160 AANARERRRMHGLNHAFDQLR 180

RESULT 2

ATH1_HUMAN STANDARD; PRT; 354 AA.
 AC Q92858;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Atonal protein homolog 1 (Helix-loop-helix protein bATH-1).
 GN ATOH1 OR ATH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97026280; PubMed=8872459;
 RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J., Zoghbi H.Y.
 RA "Evolutionary conservation of sequence and expression of the bHLH protein Atonal suggests a conserved role in neurogenesis.";
 RL Hum. Mol. Genet. 5:1207-1216(1996).
 CC -!- FUNCTION: Activates E box-dependent transcription in collaboration with E47, but the activity is completely antagonized by the negative regulator of neurogenesis HES1. May play a role in the differentiation of subsets of neural cells by activating E box-dependent transcription (By similarity).
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U61148; AAB41305.1; -;
 TRANSFAC; T04544; -;
 Genew; HGNC:797; ATOH1.
 MIM; 601461; -;
 GO; GO:0003700; P:transcription factor activity; TAS.
 GO; GO:0007417; P:central nervous system development; TAS.
 GO; GO:0006366; P:transcription from Pol II promoter; TAS.
 InterPro; IPR001092; HLH_basic.
 Pfam; PF00010; HLH; 1.
 SMART; SM00353; HLH; 1.
 PROSITE; PS00888; HLH; 1.
 Transcription regulation; Activator; DNA-binding; Nuclear protein.
 DOMAIN 29 38 POLY-PRO
 FT DNA BIND 160 171 BASIC DOMAIN
 FT DOMAIN 172 212 HELIX-LOOP-HELIX MOTIF.
 FT DOMAIN 224 228 POLY-PRO
 SQ SEQUENCE 354 AA; 38160 MW; AB12F1E917A00A8D CRC64;

Query Match 100.0%; Score 110; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
 |||||
 DB 163 AANARERRRMHGLNHAFDQLR 183

RESULT 3

AMOS DROME STANDARD; PRT; 198 AA.
 AC Q9Y0A7; Q9Y076;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Basic helix-loop-helix transcription factor Amos (Reduced olfactory organs protein) (Rough eye protein) (Absent MD neurons and olfactory sensilla protein) (Amos protein).
 DE AMOS OR ROI OR ROLO OR CG10393.
 GN Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A. AND FUNCTION.
 RP STRAIN=Oregon-R;
 RC MEDLINE=20170246; PubMed=10707973;
 RA Goulding S.E., zur Lage P., Jarman A.P.;
 RA "Amos, a proneural gene for Drosophila olfactory sense organs that is regulated by lozenge";
 RL Neuron 25:69-78(2000).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkely;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abriel J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Lasko P., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo V.G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [3]
RP FUNCTION.
RX MEDLINE=20170245; PubMed=10707972;
RA Huang M.L., Hsu C.H., Chien C.T.;
RT "The proneural gene *amos* promotes multiple dendritic neuron formation
in the *Drosophila* peripheral nervous system.";
RL Neuron 25:57-67 (2000).
CC -!- FUNCTION: Transcription factor involved in early neurogenesis.
CC Promotes multiple dendritic (MD) neuron formation. Required for
CC olfactory sensilla.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Interacts with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- DEVELOPMENTAL STAGE: During embryonic development, *amos* is
CC expressed in patches of ectodermal cells, and the expression is
CC quickly restricted to sensory organ precursors.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF166113; AAD45410.1; -;
CC DR EMBL; AE003659; AAF53678.1; -;
CC DR HSP; P25912; 1HLO.
CC DR FlyBase; FBgn0003270; *amos*.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS50888; HLH; 1.
KW Nuclear protein; Transcription regulation; Developmental protein;
KW Neurogenesis.
FT DOMAIN 111 128 SRR-RICH.
FT DNA_BIND 138 150 BASIC DOMAIN.
FT DOMAIN 151 191 HELIX-LOOP-HELIX MOTIF.
FT CONFLICT 25 25 S -> G (IN REF. 1).
SQ SEQUENCE 198 AA; 22569 MW; F1D22AF2BA1C670B CRC64;

Query Match 76.4%; Score 84; DB 1; Length 198;
Best Local Similarity 81.0%; Pred. No. 2e-06;

Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 AANARERRRMHGLNHAFDQLR 21
DB 142 AANARERRRMNSLNDAFDKLR 162
RESULT 4
ATO DROME STANDARD; PRT; 312 AA.
AC P48987; Q9VHU0;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atonal protein.
GN ATO OR CG7508.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Oregon-R;
RX MEDLINE=93313961; PubMed=8324823;
RA Jarman A.P., Grau Y., Jan L.Y., Jan Y.N.;
RT "Atonal is a proneural gene that directs chordotonal organ formation
in the *Drosophila* peripheral nervous system.";
RL Cell 73:1307-1321 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Fabios B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo V.G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
RN [3]
RP FUNCTION.

```

RX MEDLINE=94255014; PubMed=8196767;
RA Jarman A.P., Grell E.H., Ackerman L., Jan L.Y., Jan Y.N.;
RT "Atonal is the proneural gene for Drosophila photoreceptors.";
RL Nature 369:398-400(1994).
CC -!- FUNCTION: Developmental protein involved in neurogenesis. Required
CC for the formation of chordotonal organs and photoreceptors. Seems
CC to bind to E boxes. Specifically required for the photoreceptor R8
CC selection.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Forms a heterodimer with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Proneural clusters and sense organ precursors
CC of the chordotonal organs, optic furrow of the eye-antennal disk
CC and developing brain lobe.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L36646; AAA21879.1; -.
CC EMBL; AE003678; AAF54209.1; -.
CC PIR; A40708; A40708.
CC FlyBase; FBgn010433; ato.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0007173; P:EGF receptor signaling pathway; IGI.
CC GO; GO:0007605; P:hearing; IMP.
CC GO; GO:0007438; P:neocortex development; IMP.
CC GO; GO:0007422; P:peripheral nervous system development; NAS.
CC GO; GO:0045464; P:R8 cell fate specification; NAS.
CC GO; GO:0007224; P:smoothed signaling pathway; IGI.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC Neurogenesis; Differentiation; Developmental protein; Nuclear protein;
CC Transcription regulation; DNA-binding.
CC FT DNA BIND 255 267 BASIC DOMAIN.
CC FT DOMAIN 268 308 HELIX-LOOP-HELIX MOTIF.
CC FT CONFLICT 149 149 G -> A (IN REF. 1).
CC SQ SEQUENCE 312 AA; 34116 MW; 069479287438F456 CRC64;

Query Match 76.4%; Score 84; DB 1; Length 312;
Best Local Similarity 81.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAQFDQLR 21
Db 259 AANARERRRMQNLAQFDQLR 279

RESULT 5
NDP6 HUMAN
ID NDP6 HUMAN STANDARD; PRT; 337 AA.
AC Q36NR8; Q9H3H6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6) (Myo51 protein).
GN NEUROD6
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Mao Y.M., Xie Y., Zheng Z.H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

```

```

RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Tashiro H., Yanagaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Koniyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma A., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuno Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
CC with E47. May be a trans-acting factor involved in the development
CC and maintenance of the mammalian nervous system. Transactivates
CC the promoter of its own gene (By similarity).
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC in positions 300 and 307.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF063609; AAG43167.1; ALT_FRAME.
CC EMBL; AK055238; BAB70885.1; -.
CC Genew; HGNC:13804; NEUROD6.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC Neurogenesis; Differentiation; Developmental protein; Nuclear protein.
CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
CC FT DOMAIN 54 63 POLY-GLU.
CC FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 95 106 BASIC DOMAIN.
CC FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
CC SQ SEQUENCE 337 AA; 38705 MW; 6B0F4127AC0F809E CRC64;

Query Match 73.6%; Score 81; DB 1; Length 337;
Best Local Similarity 80.8%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AANARERRRMHGLNHAQFDQLR 21
Db 99 AANARERRRMHGLNLDLNL 118

RESULT 6
NDP6 MOUSE
ID NDP6 MOUSE STANDARD; PRT; 337 AA.
AC P48986;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6)
DE Neurogenic differentiation factor 6 (NeuroD6) (Atonal protein homolog
DE 2) (Helix-loop-helix protein MATH-2) (MATH2) (NEX-1 protein).
GN NEUROD6 OR ATOH2 OR ATH2 OR NEX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RA MEDLINE=95262673; PubMed=7744035;

```

RA Shimizu C., Akazawa C., Nakanishi S., Kageyama R.;
 RT "MATH-2, a mammalian helix-loop-helix factor structurally related to
 RT the product of Drosophila proneural gene atonal, is specifically
 RT expressed in the nervous system.";
 RL Eur. J. Biochem. 229:239-248(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA MEDLINE=95200803; PubMed=7545978;
 RA Bartholomae A., Nave K.-A.;
 RT "NEX-1: a novel brain-specific helix-loop-helix protein with
 RT autoregulation and sustained expression in mature cortical neurons.";
 RL Mech. Dev. 48:217-228(1994).
 CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
 CC with E47. May be a trans-acting factor involved in the development
 CC and maintenance of the mammalian nervous system. Transactivates
 CC the promoter of its own gene.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Specific to the nervous system of both embryos
 CC and adults. Highest levels in the cortical plate of the cerebrum.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D44480; BAA07923.1; -;
 DR EMBL; U29086; AAC14576.1; -;
 DR PIR; I48682; I48682.
 DR PIR; I57038; I57038.
 DR MGI; MGI:106593; Neurod6.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50888; HLH; 1.
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
 FT DOMAIN 54 63 POLY-GLU.
 FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DNA_BIND 95 106 BASIC DOMAIN.
 FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
 SQ SEQUENCE 337 AA; 38644 MW; 35C18ACD8EE1EFBA CRC64;
 Query Match 73.6%; Score 81; DB 1; Length 337;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ANARERRMHGHNHAFDQLR 21
 DB 99 ANARERNMHGLDNLRL 118
 RESULT 7
 NDFF1_XENLA
 ID NDFF1_XENLA STANDARD; PRT; 352 AA.
 AC Q91616;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 1 (Neurod1).
 GN NEUROD1 OR NEUROD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=95273957; PubMed=7754368;
 RA Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.;
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RT helix-loop-helix protein.";
 RL Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U28067; AAC59675.1; -;
 DR PIR; I51687; I51687.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50888; HLH; 1.
 KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
 KW Neurogenesis; Developmental protein; Differentiation.
 FT DOMAIN 59 78 GLU-RICH (ACIDIC)
 FT DOMAIN 88 94 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DNA_BIND 103 114 BASIC DOMAIN.
 FT DOMAIN 115 155 HELIX-LOOP-HELIX MOTIF.
 SQ SEQUENCE 352 AA; 39662 MW; 226298DB3D48233E CRC64;
 Query Match 73.6%; Score 81; DB 1; Length 352;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ANARERRMHGHNHAFDQLR 21
 DB 107 ANARERNMHGLDNLRL 126
 RESULT 8
 NDFF4_MOUSE
 ID NDFF4_MOUSE STANDARD; PRT; 330 AA.
 AC O09105;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 4 (NeuroD4) (Atonal protein homolog
 DE 3) (Helix-loop-helix protein MATH-3) (MATH3).
 GN NEUROD4 OR ATOH3 OR ATH3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=97180939; PubMed=9029157;
 RA Takebayashi K., Takahashi S., Yokota C., Tsuda H., Nakanishi S.,
 RA Asashima M., Kageyama R.;
 RT "Conversion of ectoderm into a neural fate by ATH-3, a vertebrate
 RT basic helix-loop-helix gene homologous to Drosophila proneural gene
 RT atonal.";
 RL EMBO J. 16:384-395(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/J;

RX MEDLINE=98165816; PubMed=9497361;
RA Tsuda H., Takebayashi K., Nakanishi S., Kageyama R.;
RT "Structure and promoter analysis of Math3 gene, a mouse homolog of
RT Drosophila proneural gene atonal. Neural-specific expression by dual
RT promoter elements.";
RL J. Biol. Chem. 273:6327-6333 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=223388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Expressed in the developing nervous system,
CC with high levels of expression in the brain, retina and cranial
CC ganglions.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D85845; BAA12880.1; -;
CC EMBL; AF036257; AAC15969.1; -;
CC EMBL; BC054391; AAH54391.1; -;
CC MGD; MGI:108055; Neurod4.
CC GO; GO:0001654; P:eye morphogenesis; IMP.
CC GO; GO:0007400; P:neuroblast cell fate determination; IMP.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 49 65 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 88 99 BASIC DOMAIN.
CC FT DOMAIN 100 140 ASP/GLU-RICH (ACIDIC).
CC FT DOMAIN 162 183 BASIC DOMAIN.
CC SEQUENCE 330 AA; 37133 MW; F8EC228F4EE0FF88 CRC64;

Query Match 72.7%; Score 80; DB 1; Length 330;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRRMEGLNHAFFDLR 21
||||| ||||| ||||| ||||| |||||
Db 92 ANARERTMEGLNDALDNL 111

RESULT 9
ID NDFM_CHICK STANDARD; PRT; 330 AA.
AC P79726;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurogenic differentiation factor Neurom.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Retina;
RX MEDLINE=97454246; PubMed=9310321;
RA Roztocil T., Matter-Sadzinski L., Alliod C., Ballivet M.,
RA Matter J.M.;
RT "Neurom, a neural helix-loop-helix transcription factor, defines a
RT new transition stage in neurogenesis.";
RL Development 124:3263-3272(1997).
CC -1- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC Transcriptional activator. Binds DNA on E-box consensus sequence
CC 5'-CANNTG-3'.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Developing nervous system.
CC -1- DEVELOPMENTAL STAGE: Expression in the developing nervous system
CC is transient and restricted to cells lining the ventricular zone
CC that have ceased proliferation but have not yet begun to migrate
CC into the outer layers. In retina, neurom is also transiently
CC expressed in cells as they withdraw from the mitotic cycle, but
CC persists in horizontal and bipolar neurons until full
CC differentiation. In the peripheral nervous system, its expression
CC closely follows cell proliferation.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09597; CAA70785.1; -;
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DOMAIN 49 65 GLU-RICH (ACIDIC).
CC FT DOMAIN 55 61 POLY-LYS.
CC FT DOMAIN 73 76 POLY-LYS.
CC FT DOMAIN 73 79 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 73 99 BASIC DOMAIN.
CC FT DNA BIND 88 99 BASIC DOMAIN.
CC FT DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF.
CC FT DOMAIN 162 183 LEUCINE-ZIPPER (POTENTIAL).
CC SEQUENCE 330 AA; 36707 MW; 96492F06A3C07B41 CRC64;

Query Match 72.7%; Score 80; DB 1; Length 330;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRRMEGLNHAFFDLR 21
||||| ||||| ||||| ||||| |||||
Db 92 ANARERTMEGLNDALDNL 111

RESULT 10

```

NDF4 HUMAN
ID NDF4 HUMAN STANDARD; PRT; 331 AA.
AC Q9HD90;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 4 (Neurod4).
GN NEUROD4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT THR-68.
RX MEDLINE=20527902; PubMed=11078465;
RA Horikawa Y., Horikawa Y., Cox N.J., Iwasaki N., Ogata M., Iwamoto Y.,
RA Schwitzgebel V., German M.S., Bell G.I.;
RT "beta-cell transcription factors and diabetes: no evidence for
RT diabetes-associated mutations in the gene encoding the basic
RT helix-loop-helix transcription factor neurogenic differentiation 4
RT (NEUROD4) in Japanese patients with MODY."
RL Diabetes 49:1955-1957 (2000).
CC -!- FUNCTION: Appears to mediate neuronal differentiation (By
CC similarity).
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF203901; AAF99097.1; -.
DR Genew; HGNC:13802; NEUROD4.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation;
FT DOMAIN 49 65 ASP/GLU-RICH (ACIDIC).
FT DNA_BIND 88 99 BASIC DOMAIN.
FT DOMAIN 100 140 HELIX-LOOP-HELIX MOTIF.
FT VARIANT 68 68 K -> T.
FT /FTID=VAR_012979.
SQ SEQUENCE 331 AA; 36982 MW; 56327BF2780ABF2B CRC64;

Query Match 72.7%; Score 80; DB 1; Length 331;
Best Local Similarity 80.0%; Pred. No. 1.6e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ANARERRMHGLNHAFDQLR 21
| | | | | | | | | | | | | | | | | |
Db 92 ANARERTMHGLNDALDNL 111
| | | | | | | | | | | | | | | | | |

RESULT 11
NDF1 MESAU
ID NDF1 MESAU STANDARD; PRT; 355 AA.
AC Q60430;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 1 (Neurod1) (Beta-cell E-box trans-
DE activator 2) (BETA2).
GN NEUROD1 OR NEUROD.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293222; PubMed=7774807;
RA Naya F.J., Stellrecht C.M.M., Tsai M.-J.;
RA "Tissue-specific regulation of the insulin gene by a novel basic
RT helix-loop-helix transcription factor."
RL Genes Dev. 9:1009-1019 (1995).
CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Heterodimer with E47.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Most abundant in pancreatic alpha- and beta-
CC cells, less in brain and intestine.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24679; AAA86518.1; ALT_INIT.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DOMAIN 58 77 GLU-RICH (ACIDIC).
FT DOMAIN 86 92 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DNA_BIND 101 112 BASIC DOMAIN.
FT DOMAIN 113 153 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 67 75 POLY-GLU.
FT DOMAIN 86 89 POLY-LYS.
SQ SEQUENCE 355 AA; 39763 MW; F4344DFD360226B2 CRC64;

Query Match 72.7%; Score 80; DB 1; Length 355;
Best Local Similarity 80.0%; Pred. No. 1.7e-05;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ANARERRMHGLNHAFDQLR 21
| | | | | | | | | | | | | | | | | |
Db 105 ANARERNMHGLNAALDNL 124
| | | | | | | | | | | | | | | | | |

RESULT 12
NDF1 HUMAN
ID NDF1 HUMAN STANDARD; PRT; 356 AA.
AC Q13562; Q00343; Q13340; Q96TH0; Q99455;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 1 (Neurod1) (Neurod).
GN NEUROD1 OR NEUROD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96374836; PubMed=8786144;
RA Tamimi R., Steingrimsson E., Copeland N.G., Dyer-Montgomery K.,
RA Lee J.E., Hernandez R., Jenkins N.A., Tapscott S.J.;
RT "The NEUROD gene maps to human chromosome 2q32 and mouse chromosome
RT 2."
RL Genomics 34:418-421 (1996).
RN [2]

```


RC STRAIN=White leghorn; TISSUE=Retina;
 RA MEDLINE=97454246; PubMed=9310321;
 RX Roztocil T., Matter-Sadzinski L., Alliod C., Ballivet M., Matter J.M.;
 RT "NeuroD, a neural helix-loop-helix transcription factor, defines a
 RL new transition stage in neurogenesis.";
 RN Development 124:3263-3272(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=Brain;
 RX MEDLINE=98410737; PubMed=9740021;
 RA Yan R.-T., Wang S.-Z.;
 RT "NeuroD induces photoreceptor cell overproduction in vivo and de novo
 RL generation in vitro.";
 RL J. Neurobiol. 36:485-496(1998).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC Induces photoreceptor cell overproduction in vivo and de novo
 CC generation in vitro. May play a role in photoreceptor cell
 CC production. Transcriptional activator. Binds DNA on e-box
 CC consensus sequence 5'-CANNTG-3'.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- DEVELOPMENTAL STAGE: In the spinal cord it is found in neurons
 CC that are migrating or have reached their final position. Expressed
 CC in cells located at the outer portion of the developing retinal
 CC neuroepithelium, the location where prospective photoreceptors
 CC reside.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; Y09596; CAA70784.1; -;
 CC EMBL; AF060885; AAC79425.1; -;
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
 CC Neurogenesis; Developmental protein; Differentiation.
 CC FT DOMAIN 58 77 GLU-RICH (ACIDIC).
 CC FT DOMAIN 90 96 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DNA BIND 105 116 BASIC DOMAIN.
 CC FT DOMAIN 117 157 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 69 78 POLY-GLU.
 CC FT DOMAIN 90 93 POLY-LYS.
 CC FT CONFLICT 282 282 N -> Y (IN REF. 2).
 CC FT CONFLICT 297 297 K -> N (IN REF. 2).
 CC FT CONFLICT 331 331 G -> A (IN REF. 2).
 CC SQ SEQUENCE 357 AA; 38809 MW; 770649CF9DDC54F6 CRC64;
 Query Match 72.7%; Score 80; DB 1; Length 357;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 ANARERRRMHGLNHAFDOLR 21
 IDNF1_MOUSE
 Db 109 ANARERRRMHGLNHAALDNL 128
 RESULT 14
 IDNF1_MOUSE
 AC Q60857; Q60897; PRT; 357 AA.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1) (Basic helix-loop-helix
 factor 1) (BHF-1).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GN NEUROD1 OR NEUROD.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MF1, and 129/SV;
 RC MEDLINE=95273957; PubMed=7754368;
 RX Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.;
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RL helix-loop-helix protein.";
 RL Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC Transcriptional activator. Binds to the insulin gene E-box.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein. Heterodimer with E47.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; U28068; AAC52203.1; -;
 CC EMBL; U28888; AAC52204.1; -;
 CC PIR; I49338; I49338.
 CC MGD; MGI:133708; Neurod1.
 CC InterPro; IPR001092; HLH_basic.
 CC GO; GO:0001654; P:eye morphogenesis; IMP.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS50888; HLH; 1.
 CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
 CC Neurogenesis; Developmental protein; Differentiation.
 CC FT DOMAIN 58 77 GLU-RICH (ACIDIC).
 CC FT DOMAIN 87 93 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DNA BIND 102 113 BASIC DOMAIN.
 CC FT DOMAIN 114 154 HELIX-LOOP-HELIX MOTIF.
 CC FT DOMAIN 58 64 POLY-GLU.
 CC FT DOMAIN 67 77 POLY-GLU.
 CC FT DOMAIN 87 90 POLY-LYS.
 CC SQ SEQUENCE 357 AA; 39998 MW; B6626E1315E31027 CRC64;
 Query Match 72.7%; Score 80; DB 1; Length 357;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 ANARERRRMHGLNHAFDOLR 21
 IDNF1_MOUSE
 Db 106 ANARERRRMHGLNHAALDNL 125
 RESULT 15
 IDNF1_MOUSE
 ID NF1_MOUSE
 AC Q64289;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1) (Basic helix-loop-helix
 factor 1) (BHF-1).
 GN NEUROD1 OR NEUROD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Page 10

OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RI	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Cerebellum;
RC	MEDLINE=9620182; PubMed=8660336;
RX	Kawakami H., Maruyama H., Yasunami M., Ohkubo H., Hara H., Saïda T.,
RA	Nakanishi S., Nakamura S.;
RA	"Cloning and expression of a rat brain basic helix-loop-helix
RT	factor.";
RT	Biochem. Biophys. Res. Commun. 221:199-204 (1996).
RL	[2]
RL	SEQUENCE OF 88-200 FROM N.A.
RP	STRAT-Sprague-Dawley; TISSUE=Retina;
RA	Ahmad I., Acharay H.R.;
RC	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RL	-1- FUNCTION: Acts as a differentiation factor during neurogenesis.
CC	Transcriptional activator. Binds to the insulin gene E-box.
CC	-1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC	BHLH protein. Heterodimer with E47.
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	-1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; D82075; BAAL1536.1; -;
DR	EMBL; D82074; BAA11535.1; -;
DR	EMBL; U80603; AAB38744.1; -;
DR	PIR; JC4703; JC4703.
DR	InterPro; IPR001092; HLH_basic.
DR	Pfam; PF00010; HLH; 1.
DR	SMART; SMO0353; HLH; 1.
DR	PROSITE; PS50888; HLH; 1.
KW	DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW	Neurogenesis; Developmental; Differentiation.
FT	DOMAIN 58 77 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN 87 93 BASIC DOMAIN
FT	DNA BIND 102 113 HELIX-LOOP-HELIX MOTIF.
FT	DOMAIN 114 154 POLY-GLU.
FT	DOMAIN 67 76 POLY-LYS.
FT	DOMAIN 87 90
SQ	SEQUENCE 357 AA; 40000 MW; F773637B64D3E99E CRC64;
Query Match	72.7%; Score 80; DB 1; Length 357;
Best Local Similarity	80.0%; Pred No; 1.7e-05;
Matches	16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	2 ANAFERRMHGLNHFADQLR 21
Db	106 ANAFERRMHGLNADLNLR 125

Search completed: September 21, 2004, 21:21:40
Job time : 7.624 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:04:52 ; Search time 8.12 Seconds
(without alignments)
815.995 Million cell updates/sec

Title: US-09-980-381A-70
Perfect score: 110
Sequence: 1 AANARERRRMEGLNHAFDQLR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	64	13	Q8JHA1 serinus can
2	110	100.0	161	13	Q98939 gallus gall
3	110	100.0	177	13	Q800U5 gallus gall
4	107	97.3	266	13	O42245 brachydanio
5	98	89.1	46	13	Q98904 fugu rubrip
6	91	82.7	45	5	Q94791 tribolium c
7	88	80.0	134	13	Q8JHA0 serinus can
8	88	80.0	149	11	Q922E5 mus musculus
9	88	80.0	151	13	O57598 gallus gall
10	88	80.0	152	4	Q8N100 homo sapien
11	87	79.1	138	13	O13126 xenopus lae
12	87	79.1	138	13	O13125 xenopus lae
13	85	77.3	134	13	Q9DGA9 brachydanio
14	85	77.3	134	13	Q8AW52 brachydanio
15	84	76.4	45	5	Q94790 tribolium c
16	83	75.5	202	5	Q8T360 platynereis

17	83	75.5	325	13	Q9W6C6	Q9W6C6 brachydanio
18	82	74.5	189	5	Q9XZC7	Q9XZC7 drosophila
19	82	74.5	189	5	Q9V7M2	Q9V7M2 drosophila
20	81	73.6	337	4	Q8IYR9	Q8IYR9 homo sapien
21	81	73.6	337	4	Q7Z4W9	Q7Z4W9 homo sapien
22	80	72.7	216	13	Q9W6B8	Q9W6B8 eleutheroda
23	80	72.7	255	4	Q8IWS6	Q8IWS6 homo sapien
24	80	72.7	347	13	Q9DE43	Q9DE43 brachydanio
25	80	72.7	347	13	Q8JH35	Q8JH35 brachydanio
26	80	72.7	356	4	Q9UEC8	Q9UEC8 homo sapien
27	80	72.7	357	11	Q8CEI7	Q8CEI7 mus musculus
28	79	71.8	166	4	Q96RJ6	Q96RJ6 homo sapien
29	79	71.8	168	11	Q923Z4	Q923Z4 mus musculus
30	79	71.8	336	13	Q9DDQ8	Q9DDQ8 xenopus lae
31	78	70.9	195	5	Q9VGJ5	Q9VGJ5 drosophila
32	77	70.0	151	13	Q8AX95	Q8AX95 fugu rubrip
33	77	70.0	232	5	Q9VJCI	Q9VJCI drosophila
34	77	70.0	238	13	Q7T301	Q7T301 brachydanio
35	77	70.0	350	13	O42202	O42202 brachydanio
36	77	70.0	355	11	Q8C6A8	Q8C6A8 mus musculus
37	77	70.0	355	11	Q9JL05	Q9JL05 mus musculus
38	77	70.0	381	4	Q8NFK8	Q8NFK8 homo sapien
39	76	69.1	316	13	Q9W6C7	Q9W6C7 brachydanio
40	73	66.4	220	13	Q8JIS0	Q8JIS0 cynops pyrr
41	73	66.4	223	11	Q8BGW3	Q8BGW3 mus musculus
42	73	66.4	225	4	Q8NDY6	Q8NDY6 homo sapien
43	73	66.4	255	5	Q9NB27	Q9NB27 branchiosto
44	73	66.4	381	4	Q9UQC6	Q9UQC6 homo sapien
45	73	65.4	382	4	Q8TBI7	Q8TBI7 homo sapien

ALIGNMENTS

RESULT 1

Q8JHA1 PRELIMINARY; PRT; 64 AA.

AC Q8JHA1;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE BHLH transcription factor ath1 (Fragment).
OS Serinus canaria (Canary).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC Fringillidae; Carduelinae; Serinus.
CX NCBI_TaxID=9135;
RN [1]
RP SEQUENCE FROM N.A.
RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,
RA Ryan A.F.;
RT "Expression of BHLH and class IV POU-domain transcription factors in
RT the chicken and canary inner ear";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF526421; AAM89247.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SMO0353; HLH; 1.
DR PROSITE; PS50888; HLH_2; 1.
FT NON_TER 1
FT NON_TER 64
SQ SEQUENCE 64 AA; 7439 MW; BB691E056670CAB4 CRC64;

Query Match 100.0%; Score 110; DB 13; Length 64;

Best Local Similarity 100.0%; Pred. No. 4.4e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AANARERRRMEGLNHAFDQLR 21

Db 11 AANARERRRMEGLNHAFDQLR 31

```

Query Match      100.0%; Score 110; DB 13; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   |||||
Db 12 AANARERRRMHGLNHAFDQLR 32
   |||||

RESULT 4
O42245 PRELIMINARY; PRT; 266 AA.
AC O42245;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Atonal homologue-1.
GN ATOH1 OR ZATH-1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.-H., Bae Y.-K., Yamanaka Y., Yamashita S., Shimizu T., Fujii R.,
RA Hibi M., Hirano T.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF024536; AAB82272.1; -.
DR ZFIN; ZDB-GENE-990415-17; atohl.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00868; HLH_2; 1.
DR NON TER 161
FT SEQUENCE 161 AA; 16652 MW; 56A92494B52FEFB8 CRC64;
SQ

Query Match      97.3%; Score 107; DB 13; Length 266;
Best Local Similarity 95.2%; Pred. No. 5.6e-09;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   |||||
Db 122 AANARERRRMHGLNHAFDQLR 142
   |||||

RESULT 5
Q98904 PRELIMINARY; PRT; 46 AA.
AC Q98904;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FATH1 (Fragment).
GN FATH1.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Notoleosteii;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
RA Zoghbi H.Y.;
RL "Evolutionary conservation of sequence and expression of the bHLH
RL protein Atonal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U61150; AAB41303.1; -.

Query Match      100.0%; Score 110; DB 13; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
   |||||
Db 89 AANARERRRMHGLNHAFDQLR 109
   |||||

RESULT 3
Q80005 PRELIMINARY; PRT; 177 AA.
AC Q80005;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CATH1 (Fragment).
GN CATH1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert P.J., Timmer J.R., Helms A.W., Nakada Y., Hunsaker T.L.,
RA Johnson J.E.;
RL "Control of Math1 Expression by Interactions of Zic1 and Ebf2 with the
RL Math1 Enhancer.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467292; AAO59913.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00868; HLH_2; 1.
DR NON TER 1
FT SEQUENCE 177 AA; 19170 MW; FD44269C5994F5E9 CRC64;
SQ
```

DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS50888; HLH 2; 1.
 FT NON_TER 1
 FT NON_TER 46
 SQ SEQUENCE 46 AA; 5422 MW; 4F88FFD41319483C CRC64;

Query Match 89.1%; Score 98; DB 13; Length 46;
 Best Local Similarity 90.5%; Pred. No. 2.6e-08;
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMEGLNKAFLDLR 21

RESULT 6

Q94791 ID Q94791 PRELIMINARY; PRT; 45 AA.

AC Q94791;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE TATH2 (Fragment).

GN TATH2.
 OS Tribolium castaneum (Red flour beetle).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Tenebrionidae; Tribolium.

OX NCBI_TaxID=7070;

RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=97026280; PubMed=9872459;

RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
 Zoghbi H.F.;

RT Evolutionary conservation of sequence and expression of the bHLH

protein Atonal suggests a conserved role in neurogenesis.";

RL Hum. Mol. Genet. 5:1207-1216(1996).

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

TRANSCRIPTION FACTORS.

CC EMBL; U61152; AAB41307.1; -.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS50888; HLH 2; 1.

FT NON_TER 1

FT NON_TER 45

SQ SEQUENCE 45 AA; 5263 MW; 730A6F8921B82E23 CRC64;

Query Match 82.7%; Score 91; DB 5; Length 45;
 Best Local Similarity 85.7%; Pred. No. 3.3e-07;
 Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||
 Db 1 AANARERRMEGLNEAFDLR 21

RESULT 7

Q8UHAO ID Q8UHAO PRELIMINARY; PRT; 134 AA.

AC Q8UHAO;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE BHLH transcription factor at5 (Fragment).

OS Serinus canaria (Canary).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;

OC Fringillidae; Carduelinae; Serinus.

OX NCBI_TaxID=9135;

RN [1]

RP SEQUENCE FROM N.A.
 RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,
 RA Ryan A.F.;

RT "Expression of bHLH and class IV POU-domain transcription factors in

the chicken and canary inner ear.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

TRANSCRIPTION FACTORS.

DR EMBL; AF526422; AAM89248.1; -.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS50888; HLH 2; 1.

FT NON_TER 134

FT NON_TER 134

SQ SEQUENCE 134 AA; 15448 MW; 0016C7C6807F0D6C CRC64;

Query Match 80.0%; Score 88; DB 13; Length 134;
 Best Local Similarity 85.7%; Pred. No. 3.1e-06;
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||

Db 43 AANARERRMQGLNTAFDLR 63

RESULT 8

Q9Z2E5 ID Q9Z2E5 PRELIMINARY; PRT; 149 AA.

AC Q9Z2E5;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Transcription factor MATH5 (ATOH7).

GN ATOH7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=129SV/J; TISSUE=Retina;

RX MEDLINE=99026071; PubMed=9806930;

RA Brown N.L., Kanekar S., Vetter M.L., Tucker P.K., Gemza D.L.,

Glaser T.;

RT "Math5 encodes a murine basic helix-loop-helix transcription factor

expressed during early stages of retinal neurogenesis.";

RL Development 125:4821-4833(1998).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=129S4;

RX MEDLINE=1886645; PubMed=11889557;

RA Brown N.L., Dagenais S.L., Chen C.M., Glaser T.;

RT "Molecular characterization and mapping of ATOH7, a human atonal

homolog with a predicted role in retinal ganglion cell development.";

RL Mamm. Genome 13:95-101(2002).

DR EMBL; AF071223; AAC68868.1; -.

DR EMBL; AF418923; AAL11912.1; -.

DR MGD; MGI:135553; Atoh7.

DR GO; GO:0009649; P:entrainment of circadian clock; IMP.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR PROSITE; PS50888; HLH 2; 1.

SQ SEQUENCE 149 AA; 16569 MW; DF78A25190B5FD3A CRC64;

Query Match 80.0%; Score 88; DB 11; Length 149;
 Best Local Similarity 85.7%; Pred. No. 3.4e-06;
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRMEGLNHAFDQLR 21
 |||||

Db 45 AANARERRMQGLNTAFDLR 65

```

RESULT 9
O57598      PRELIMINARY;          PRT; 151 AA.
ID   O57598;
AC   O57598;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Atonal transcription factor homologue.
GN   ATH5.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21064448; PubMed=11124117;
RA   Matter-Sadzinski L., Matter J.M., Ong M.T., Hernandez J., Ballivet M.;
RT   "Specification of neurotransmitter receptor identity in developing
RT   retina: the chick ATH5 promoter integrates the positive and negative
RT   effects of several bHLH proteins.";
RL   Development 128:217-231 (2001).
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC   TRANSCRIPTION FACTORS.
DR   EMBL; AJ001178; CAA04572.1; -.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH 2; 1.
FT   CHAIN 2 151
SQ   SEQUENCE 151 AA; 17540 MW; 1D42F9D225A5319 CRC64;

Query Match      80.0%; Score 88; DB 13; Length 151;
Best Local Similarity 85.7%; Pred. No. 3.4e-06;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY   1 AANARERRRMHGLNHAFDQLR 21
    |||||
DB   43 AANARERRRMQGLNTAFDRLR 63
    |||||

RESULT 10
O8N100      PRELIMINARY;          PRT; 152 AA.
ID   O8N100;
AC   O8N100;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Atonal homolog 7 (ATOH7).
GN   ATOH7.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21886645; PubMed=11889557;
RA   Brown N.L., Dagenais S.L., Chen C.M., Glaser T.;
RT   "Molecular characterization and mapping of ATOH7, a human atonal
RT   homolog with a predicted role in retinal ganglion cell development.";
RL   Mamm. Genome 13:95-101 (2002).
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC   TRANSCRIPTION FACTORS.
DR   EMBL; BC032621; AAB32621.1; -.
DR   EMBL; AF418922; AAL11911.1; -.
DR   Genew; HGNC:113907; ATOH7.
DR   InterPro; IPR001092; HLH_basic.

```

```

DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH 2; 1.
SQ   SEQUENCE 152 AA; 16871 MW; 9E93E9E60E1697C5 CRC64;

Query Match      80.0%; Score 88; DB 4; Length 152;
Best Local Similarity 85.7%; Pred. No. 3.5e-06;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY   1 AANARERRRMHGLNHAFDQLR 21
    |||||
DB   44 AANARERRRMQGLNTAFDRLR 64
    |||||

RESULT 11
O13126      PRELIMINARY;          PRT; 138 AA.
ID   O13126;
AC   O13126;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Atonal homolog 5b.
GN   XATH5B.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Head;
RA   Kanekar S., Perron M., Harris W.A., Jan L.Y., Jan Y.N., Vetter M.L.;
RT   "Xath5 participates in a network of bHLH genes in the developing
RT   Xenopus retina.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC   TRANSCRIPTION FACTORS.
DR   EMBL; U93171; AAB58669.1; -.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH 2; 1.
SQ   SEQUENCE 138 AA; 16028 MW; A72A5DD9BE118EDE CRC64;

Query Match      79.1%; Score 87; DB 13; Length 138;
Best Local Similarity 85.7%; Pred. No. 4.5e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY   1 AANARERRRMHGLNHAFDQLR 21
    |||||
DB   37 AANARERRRMQGLNTAFDLSL 57
    |||||

RESULT 12
O13125      PRELIMINARY;          PRT; 138 AA.
ID   O13125;
AC   O13125;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Atonal homolog 5a.
GN   XATH5A.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Head;
RA   Kanekar S., Perron M., Harris W.A., Jan L.Y., Jan Y.N., Vetter M.L.;
RT   "Xath5 participates in a network of bHLH genes in the developing
RT   Xenopus retina.";

```

```
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U93170; AAB58668.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 138 AA; 16212 MW; 25363E84A6031730 CRC64;

Query Match          79.1%; Score 87; DB 13; Length 138;
Best Local Similarity 85.7%; Pred. No. 4.5e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AANARERRRMHGLNHAFDQLR 21
Db 37 AANARERRRMQGLNTAFDSL 57

RESULT 13
Q9DGA9 ID Q9DGA9 PRELIMINARY; PRT; 134 AA.
AC Q9DGA9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Atonal homolog 5.
GN ATOH7 OR ZATH5.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20439479; PubMed=10985346;
RA Masai I., Stemple D.L., Okamoto H., Wilson S.W.;
RT "Midline signals regulate retinal neurogenesis in zebrafish.";
RL Neuron 27:251-263(2000).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AB049457; BAB15953.1; -.
DR HSPG; P25912; IHLO.
DR ZFIN; ZDB-GENE-000926-1; atoh7.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 134 AA; 15591 MW; 3207D8CFD8972496 CRC64;

Query Match          77.3%; Score 85; DB 13; Length 134;
Best Local Similarity 81.0%; Pred. No. 9.2e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRRMHGLNHAFDQLR 21
Db 32 AANARERRRMQGLNTAFDRLR 52

RESULT 14
Q8AW52 ID Q8AW52 PRELIMINARY; PRT; 134 AA.
AC Q8AW52;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SI:PACKTR2.1 (Atonal homolog 7).
GN ATOH7.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL627094; CAD52125.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
SQ SEQUENCE 134 AA; 15621 MW; 3207DBD0E5972496 CRC64;

Query Match          77.3%; Score 85; DB 13; Length 134;
Best Local Similarity 81.0%; Pred. No. 9.2e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRRMHGLNHAFDQLR 21
Db 32 AANARERRRMQGLNTAFDRLR 52

RESULT 15
Q94790 ID Q94790 PRELIMINARY; PRT; 45 AA.
AC Q94790;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE TATH1 (Fragment).
GN TATH1.
OS Tribolium castaneum (Red flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tribolium.
OX NCBI_TaxID=7070;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97026280; PubMed=8872459;
RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
RA Zoghbi H.Y.;
RT "Evolutionary conservation of sequence and expression of the bHLH
RT protein Atonal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U61151; AAB41306.1; -.
DR HSPG; P25912; IHLO.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH 2; 1.
FT NON_TER 1
FT NON_TER 45
SQ SEQUENCE 45 AA; 5256 MW; 633ACFB9E1FE5F0A CRC64;

Query Match          76.4%; Score 84; DB 5; Length 45;
Best Local Similarity 81.0%; Pred. No. 4.4e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AANARERRRMHGLNHAFDQLR 21
Db 1 AANARERRRMNSLNDADFRLR 21

Search completed: September 21, 2004, 21:24:12
Job time : 9.12 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:22:55 ; Search time 0.001 Seconds
(without alignments)
10.710 Million cell updates/sec

Title: us-10-084-555a-115
Perfect score: 21
Sequence: 1 ttgtgtggggagtatttgagt 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	6.4	30.5	255	1	us-10-084-555a-8
2	5	23.8	255	1	us-10-084-555a-8

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 30.5%; Score 6.4; DB 1; Length 255;
Best Local Similarity 62.5%; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TGTGGGGAGTTATTGA 19
||| ||||| |||
Db 200 TGCAGGAGAGAGAGGA 185

RESULT 2
us-10-084-555a-8

Query Match 23.8%; Score 5; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TGAGT 21
||| |||||
Db 22 TGAGT 26

Search completed: September 22, 2004, 12:22:55
Job time : 0.001 secs

Search completed: September 22, 2004, 12:23:44
Job time : 0.001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:24:40 ; Search time 0.001 Seconds
(without alignments)
9.180 Million cell updates/sec

Title: us-10-084-555a-117
Perfect score: 18
Sequence: 1 tgggggagttatcgagc 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	7.6	42.2	255	1 us-10-084-555a-8	
2	5	27.8	255	1 us-10-084-555a-8	

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 42.2% ; Score 7.6; DB 1; Length 255;
Best Local Similarity 71.4% ; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GGGAGTTATCGAGC 18
||| |||||
Db 113 GGGGATCGTCGAGC 100

RESULT 2
us-10-084-555a-8

Query Match 27.8% ; Score 5; DB 1; Length 255;
Best Local Similarity 61.5% ; Pred. No. 0;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 GGAGTTATCGAGC 18
||| |||||
Db 79 GGAGCATTCTGC 91

Search completed: September 22, 2004, 12:24:41
Job time : 1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2004, 12:25:32 ; Search time 0.001 Seconds
(without alignments)
9.690 Million cell updates/sec

Title: us-10-084-555a-118
Perfect score: 19
Sequence: 1 gccttcgcgaaaaaatcg 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 255 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 2000 summaries

Database : us-10-084-555a-8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	7.6	40.0	255	1	us-10-084-555a-8
2	5.8	30.5	255	1	us-10-084-555a-8

ALIGNMENTS

RESULT 1
us-10-084-555a-8/c

Query Match 40.0%; Score 7.6; DB 1; Length 255;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCCTTCGCGAAAAA 14
||| ||| ||| |||
Db 168 GCCTGGGGCGAGAA 155

RESULT 2
us-10-084-555a-8

Query Match 30.5%; Score 5.8; DB 1; Length 255;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCTTCGCG 9
||| ||| ||| |||
Db 28 GCCTCCACG 36

Search completed: September 22, 2004, 12:25:32
Job time : 0.001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:34:21 ; Search time 9.912 Seconds
(without alignments)
598.618 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRMEGLNHPDQLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04: *
1: Genesep1980s: *
2: Genesep1990s: *
3: Genesep2000s: *
4: Genesep2001s: *
5: Genesep2002s: *
6: Genesep2003as: *
7: Genesep2003bs: *
8: Genesep2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	110	100.0	21	4	AAB60379	Aab60379 Atonal-as
2	110	100.0	21	6	AAO30950	Aao30950 Transcrip
3	110	100.0	161	4	AAB60376	Aab60376 Chicken a
4	110	100.0	161	6	AAO30946	Aao30946 Chicken a
5	110	100.0	259	6	AAO30949	Aao30949 Frog aton
6	110	100.0	351	4	AAB60353	Aab60353 Mouse ato
7	110	100.0	351	4	AAB60369	Aab60369 Mouse ato
8	110	100.0	351	6	AAO30939	Aao30939 Mouse ato
9	110	100.0	351	6	AAO30924	Aao30924 Mouse ato
10	110	100.0	354	4	AAB60375	Aab60375 Human ato
11	110	100.0	354	4	AAB60349	Aab60349 Human ato
12	110	100.0	354	6	AAO30945	Aao30945 Human ato
13	110	100.0	354	6	AAO30920	Aao30920 Human ato
14	107	97.3	266	4	AAB60366	Aab60366 Zebrafish
15	107	97.3	266	6	AAO30936	Aao30936 Zebra fis
16	98	89.1	46	4	AAB60377	Aab60377 T. rubrip
17	98	89.1	46	6	AAO30947	Aao30947 Puffer fi
18	91	82.7	45	4	AAB60373	Aab60373 T. casten
19	91	82.7	45	6	AAO30943	Aao30943 Beetle at
20	88	80.0	149	4	AAB60358	Aab60358 Mouse ato
21	88	80.0	149	6	AAO30928	Aao30928 Mouse ato
22	88	80.0	151	4	AAB60364	Aab60364 Chicken a
23	88	80.0	151	6	AAO30934	Aao30934 Chicken a
24	87	79.1	138	4	AAB60372	Aab60372 Xenopus l
25	87	79.1	138	4	AAB60371	Aab60371 Xenopus l

ALIGNMENTS

RESULT 1

AAB60379
ID AAB60379 standard; peptide; 21 AA.

XX AAB60379;

DT 24-APR-2001 (first entry)

DE Atonal-associated peptide fragment.

XX

KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Unidentified.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
PT of its homologs or orthologs, for the treatment of e.g. deafness,
PT osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page 33; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
XX amino acid sequence, or any of its homologues or orthologues as
XX therapeutic agents for the treatment of deafness, partial hearing loss,
XX vestibular effects due to damage or loss of inner hair cells,
XX osteoarthritis and abnormal cell proliferation. The invention also
XX encompasses methods of screening for compounds which affect the
XX expression of an atonal-associated nucleic acid sequence in an animal,
XX and a transgenic animal in which an allele of a native atonal-associated
XX gene is replaced by a heterologous nucleic acid sequence, thus

Aao30942 Frog aton
Aao30941 Frog aton
Aab60374 T. casten
Aao30944 Beetle at
Abb71919 Drosophil
Abb63330 Drosophil
Aab60378 Drosophil
Aao30948 Drosophil
Abb63483 Drosophil
Aab60355 Drosophil
Aao30925 Drosophil
Aam41883 Human pol
Aab12503 Human MAT
Aag66722 Human ato
Abb50169 Human tra
Aab60352 Mouse ato
Aab60368 Mouse ato
Aao30938 Mouse ato
Aao30923 Mouse ato
Aar77505 Frog neur

26 87 79.1 138 6 AAO30942
27 87 79.1 138 6 AAO30941
28 84 76.4 45 4 AAB60374
29 84 76.4 45 6 AAO30944
30 84 76.4 198 4 ABB71919
31 84 76.4 312 4 ABB63330
32 84 76.4 312 4 AAB60378
33 84 76.4 312 6 AAO30948
34 82 74.5 189 4 ABB63483
35 82 74.5 189 4 AAB60355
36 82 74.5 189 6 AAO30925
37 81 73.6 245 4 AAM41883
38 81 73.6 337 3 AAB12503
39 81 73.6 337 4 AAG66722
40 81 73.6 337 4 ABB50169
41 81 73.6 337 4 AAB60352
42 81 73.6 337 4 AAB60368
43 81 73.6 337 6 AAO30938
44 81 73.6 337 6 AAO30923
45 81 73.6 352 2 AAR77505

CC inactivating the atonal-associated allele. The nucleic acids or proteins
CC may be used in a method of treating an animal for hearing impairment,
CC joint disease, balance disorders, abnormal cell proliferation, or other
CC disease related to loss of a functional atonal-associated nucleic acid or
CC protein. They may particularly be used to treat an animal with a
CC deficiency in cerebellar granule neurons or their precursors, and may
CC also be used in promoting mechanoreceptive cell growth and generating
CC hair cells. The present sequence represents an atonal-associated amino
CC acid sequence referred to in the invention
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 110; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.1e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AANARERRMHGLNHAFDQLR 21
Dd 1 AANARERRMHGLNHAFDQLR 21
|||||

RESULT 2

AAO30950
ID AAO30950 standard; peptide; 21 AA.

AC AAO30950;

XX 22-SEP-2003 (first entry)

DE Transcription factor peptide used in the invention.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
KW necrotizing enterocolitis; hearing impairment; abnormal proliferation;
KW joint disease; gene therapy; imbalance disorder; transcription factor.

OS Unidentified.

XX WO2003047532-A2.

PN 12-JUN-2003.

XX 03-DEC-2002; 2002WO-US041458.

PF 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

PA Zoghbi HY, Qi Y;

XX WPI; 2003-505253/47.

XX New compositions comprising at least one stem cell which is up- or down-
PT regulated for expression of an atonal-associated sequence, and at least
PT one regulatory factor, useful for treating a gastrointestinal condition,
PT e.g. cancer.

XX Disclosure; Page 27; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
CC therapeutic use of an atonal-associated sequence. The composition
CC comprises of a stem cell which is up- or down-regulated for expression of
CC an atonal-associated sequence and a regulatory factor. Atonal-associated
CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
CC condition including cancer, damaged intestinal tissue, inflammatory bowel
CC disease, irritable bowel syndrome, infection or necrotizing enterocolitis.
CC They are also useful for promoting mechanoreceptive cell growth, for
CC generating hair cells, for treating hearing impairment or imbalance
CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
CC disease that is a result or loss of functional atonal-associated
CC sequences. Atonal-associated sequences are also used in gene therapy. The
CC present sequence is a transcription factor peptide used to illustrate the
CC method of the invention

XX
SQ

Sequence 21 AA;

Query Match 100.0%; Score 110; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 9.1e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AANARERRMHGLNHAFDQLR 21
Dd 1 AANARERRMHGLNHAFDQLR 21
|||||

RESULT 3

AAAB60376
ID AAB60376 standard; protein; 161 AA.

XX AAB60376;

XX 24-APR-2001 (first entry)

XX Chicken atonal homologue 1 (Cath1) protein, SEQ ID NO:60.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Gallus gallus.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27284.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any
PT of its homologs or orthologs, for the treatment of e.g. deafness,
PT osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or
CC amino acid sequence, or any of its homologues or orthologues as
CC therapeutic agents for the treatment of deafness, partial hearing loss,
CC vestibular effects due to damage or loss of inner hair cells,
CC osteoarthritis and abnormal cell proliferation. The invention also
CC encompasses methods of screening for compounds which affect the
CC expression of an atonal-associated nucleic acid sequence in an animal,
CC and a transgenic animal in which an allele of a native atonal-associated
CC gene is replaced by a heterologous nucleic acid sequence, thus
CC inactivating the atonal-associated allele. The nucleic acids or proteins
CC may be used in a method of treating an animal for hearing impairment,
CC joint disease, balance disorders, abnormal cell proliferation, or other
CC disease related to loss of a functional atonal-associated nucleic acid or
CC protein. They may particularly be used to treat an animal with a
CC deficiency in cerebellar granule neurons or their precursors, and may
CC also be used in promoting mechanoreceptive cell growth and generating
CC hair cells. The present sequence represents an atonal-associated amino
CC acid sequence referred to in the invention. Note: The present sequence is
CC not shown in the specification, but was obtained from GenBank

XX Sequence 161 AA;

Query Match	100.0%;	Score 110;	DB 4;	Length 161;
Best Local Similarity	100.0%;	Pred. No. 8.6e-11;		
Matches	21;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	AAARERRRMEGLNHAFDQLR	21
DB	89	AAARERRRMEGLNHAFDQLR	109

RESULT 4
IAO30946
ID AAO30946 standard; protein; 161 AA.
XX AAO30946;
AC
XX
DT 22-SEP-2003 (first entry)
DE Chicken atonal homologue 1 (Cath1) protein #4.
XX
KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
XX Cath1; chicken.
XX
OS Gallus sp.
XX
PN WO2003047532-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US041458.
XX
PR 05-DEC-2001; 2001US-00004717.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PI Zoghbi HY, Qi Y;
XX
DR WPI; 2003-505253/47.
DR N-PSDB; AAL62231.
XX
PT New compositions comprising at least one stem cell which is up- or down-
PT regulated for expression of an atonal-associated sequence, and at least
PT one regulatory factor, useful for treating a gastrointestinal condition,
PT e.g. cancer.
XX
PS Disclosure; Page 152-153; 157pp; English.
XX
CC The invention relates to pharmaceutical compositions and methods for the
CC therapeutic use of an atonal-associated sequence. The composition
CC comprises of a stem cell which is up- or down-regulated for expression of
CC an atonal-associated sequence and a regulatory factor. Atonal-associated
CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
CC condition including cancer, damaged intestinal tissue, inflammatory bowel
CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
CC They are also useful for promoting mechanoreceptive cell growth, for
CC generating hair cells, for treating hearing impairment or imbalance
CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
CC disease that is a result of loss of functional atonal-associated
CC sequences. Atonal-associated sequences are also used in gene therapy. The
CC present sequence is chicken atonal homologue 1 (Cath1) protein. This
CC sequence is used to illustrate the method of the invention
XX
SQ Sequence 161 AA;

Query Match	100.0%;	Score 110;	DB 6;	Length 161;
Best Local Similarity	100.0%;	Pred. No. 8.6e-11;		
Matches	21;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	AAARERRRMEGLNHAFDQLR	21

PT New compositions comprising at least one stem cell which is up- or down-regulated for expression of an atonal-associated sequence, and at least one regulatory factor, useful for treating a gastrointestinal condition, e.g. cancer.

PS Disclosure; Page 155-156; 157pp; English.

XX
CC The invention relates to pharmaceutical compositions and methods for the therapeutic use of an atonal-associated sequence. The composition comprises of a stem cell which is up- or down-regulated for expression of an atonal-associated sequence and a regulatory factor. Atonal-associated amino acid or nucleic acid sequence is used to treat a gastrointestinal condition including cancer, damaged intestinal tissue, inflammatory bowel disease, irritable bowel syndrome, infection or necrotising enterocolitis. They are also useful for promoting mechanoreceptive cell growth, for generating hair cells, for treating hearing impairment or imbalance disorder, joint disease, abnormal proliferation (e.g. cancer), or a disease that is a result of loss of functional atonal-associated sequences. Atonal-associated sequences are also used in gene therapy. The present sequence is frog atonal homologue 1 (Xathl) protein. This sequence is used to illustrate the method of the invention

XX Sequence 259 AA;

Query Match 100.0%; Score 110; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
|||||
Db 115 AANARERRMHGLNHAFDQLR 135

RESULT 6
AAB60353
ID AAB60353 standard; protein; 351 AA.

XX

AC AAB60353;

XX 24-APR-2001 (first entry)

XX Mouse atonal homologue 1 (ATOH1, Math1) protein, SEQ ID NO:11.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Mus musculus.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27257.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologs or orthologs, for the treatment of e.g. deafness, osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX

CC The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment, joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is not shown in the specification, but was obtained from GenBank

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
|||||
Db 160 AANARERRMHGLNHAFDQLR 180

RESULT 7

AAB60369

ID AAB60369 standard; protein; 351 AA.

XX AAB60369;

XX 24-APR-2001 (first entry)

XX Mouse atonal homologue 1 (Math-1) protein, SEQ ID NO:46.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX Mus musculus.

XX WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27277.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologs or orthologs, for the treatment of e.g. deafness, osteoarthritis and abnormal cell proliferation.

XX Disclosure; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as

therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment, joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is not shown in the specification, but was obtained from GenBank

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 4; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRHGLNHAFDQLR 21
 |||||
 Db 160 AANARERRRHGLNHAFDQLR 180

RESULT 8
 AAO30939
 ID AAO30939 standard; protein; 351 AA.

XX AAO30939;

DT 22-SEP-2003 (first entry)

DE Mouse atonal homologue 1 (Math1) protein #13.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.

XX Mus musculus.

XX WO2003047532-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-US041458.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

DR WPI; 2003-505253/47.

DR N-PSDB; AAL62224.

XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.

XX Disclosure; Page 144-145; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated

CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is mouse atonal homologue 1 (Math1) protein. This
 CC sequence is used to illustrate the method of the invention

XX Sequence 351 AA;

Query Match 100.0%; Score 110; DB 6; Length 351;

Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRHGLNHAFDQLR 21
 |||||
 Db 160 AANARERRRHGLNHAFDQLR 180

RESULT 9
 AAO30924

ID AAO30924 standard; protein; 351 AA.

XX AAO30924;

DT 22-SEP-2003 (first entry)

DE Mouse atonal homologue 1 (Math1) protein #4.

XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.

XX Mus musculus.

XX WO2003047532-A2.

XX 12-JUN-2003.

PF 03-DEC-2002; 2002WO-US041458.

XX 05-DEC-2001; 2001US-00004717.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Qi Y;

DR WPI; 2003-505253/47.

DR N-PSDB; AAL62204.

XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.

XX Disclosure; Page 122-123; 157pp; English.

XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated

SQ Sequence 354 AA;
 Query Match 100.0%; Score 110; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANARERRMHGHNHAFDQLR 21
 |||||

Db 163 AANARERRMHGHNHAFDQLR 183
 |||||

RESULT 12
 AAO30945
 ID AAO30945 standard; protein; 354 AA.
 AC AAO30945;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hath1) protein #2.
 XX
 KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hath1; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2003047532-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-US041458.
 XX
 PR 05-DEC-2001; 2001US-00004717.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Qi Y;
 XX
 DR WPI; 2003-505253/47.
 DR N-PSDB; AAL62230.
 XX
 PT New compositions comprising at least one stem cell which is up- or down-
 regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 PS Claim 33; Page 151-152; 157pp; English.
 XX
 CC The invention relates to pharmaceutical compositions and methods for the
 therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result of loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hath1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 SQ Sequence 354 AA;
 Query Match 100.0%; Score 110; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANARERRMHGHNHAFDQLR 21
 |||||

Db 163 AANARERRMHGHNHAFDQLR 183
 |||||

RESULT 14
 AAB60366
 ID AAB60366 standard; protein; 266 AA.

```

XX AC AAB60366;
XX DT 24-APR-2001 (first entry)
XX DE
XX DE Zebrafish atonal homologue-1 (zath-1) protein, SEQ ID NO:40.
XX KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
XX KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
XX KW cellular proliferation; cerebellar granule neuron; gene therapy;
XX KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
XX KW transgenic animal.
XX OS
XX OS Danio rerio.
XX PN WO200073764-A2.
XX XX
XX PD 07-DEC-2000.
XX PF
XX PF 01-JUN-2000; 2000WO-US015410.
XX PR
XX PR 01-JUN-1999; 99US-0137060P.
XX PR 19-JAN-2000; 2000US-0176933P.
XX XX
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX XX
XX PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
XX XX
XX DR WPI; 2001-0322190/04.
XX DR N-PSDB; AAF27274.
XX XX
XX PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
XX PT of its homologs or orthologs, for the treatment of e.g. deafness,
XX PT osteoarthritis and abnormal cell proliferation.
XX XX
XX PS Disclosure; Page; 142pp; English.
XX XX
XX CC The invention relates to the use of atonal-associated nucleic acid or
XX CC amino acid sequence, or any of its homologues or orthologues as
XX CC therapeutic agents for the treatment of deafness, partial hearing loss,
XX CC vestibular effects due to damage or loss of inner hair cells,
XX CC osteoarthritis and abnormal cell proliferation. The invention also
XX CC encompasses methods of screening for compounds which affect the
XX CC expression of an atonal-associated nucleic acid sequence in an animal,
XX CC and a transgenic animal in which an allele of a native atonal-associated
XX CC gene is replaced by a heterologous nucleic acid sequence, thus
XX CC inactivating the atonal-associated allele. The nucleic acids or proteins
XX CC may be used in a method of treating an animal for hearing impairment,
XX CC joint disease, balance disorders, abnormal cell proliferation, or other
XX CC disease related to loss of a functional atonal-associated nucleic acid or
XX CC protein. They may particularly be used to treat an animal with a
XX CC deficiency in cerebellar granule neurons or their precursors, and may
XX CC also be used in promoting mechanoreceptive cell growth and generating
XX CC hair cells. The present sequence represents an atonal-associated amino
XX CC acid sequence referred to in the invention. Note: The present sequence is
XX CC not shown in the specification, but was obtained from GenBank
XX XX
XX SQ Sequence 266 AA;
XX
Query Match 97.3%; Score 107; DB 4; Length 266;
Best Local Similarity 95.2%; Pred. No. 4.9e-10;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
Db 122 AANARERRRMHGLNHAFDEL 142

RESULT 15
AAO30936
ID AAO30936 standard; protein; 266 AA.
XX
AC AAO30936;

```

```

XX DT 23-OCT-2003 (revised)
XX DT 22-SEP-2003 (first entry)
XX XX
XX DE Zebra fish atonal homologue 1 protein #1.
XX KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
XX KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
XX KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
XX KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
XX KW zebra fish.
XX OS
XX OS Danio rerio.
XX PN WO2003047532-A2.
XX XX
XX PD 12-JUN-2003.
XX PF
XX PF 03-DEC-2002; 2002WO-US041458.
XX PR
XX PR 05-DEC-2001; 2001US-00004717.
XX XX
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX XX
XX PI Zoghbi HY, Qi Y;
XX XX
XX DR WPI; 2003-505253/47.
XX DR N-PSDB; AAL62221.
XX XX
XX PT New compositions comprising at least one stem cell which is up- or down-
XX PT regulated for expression of an atonal-associated sequence, and at least
XX PT one regulatory factor, useful for treating a gastrointestinal condition,
XX PT e.g. cancer.
XX XX
XX PS Disclosure; Page 139-140; 157pp; English.
XX XX
XX CC The invention relates to pharmaceutical compositions and methods for the
XX CC therapeutic use of an atonal-associated sequence. The composition
XX CC comprises of a stem cell which is up- or down-regulated for expression of
XX CC an atonal-associated sequence and a regulatory factor. Atonal-associated
XX CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
XX CC condition including cancer, damaged intestinal tissue, inflammatory bowel
XX CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
XX CC They are also useful for promoting mechanoreceptive cell growth, for
XX CC generating hair cells for treating hearing impairment or imbalance
XX CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
XX CC disease that is a result or loss of functional atonal-associated
XX CC sequences. Atonal-associated sequences are also used in gene therapy. The
XX CC present sequence is zebra fish atonal homologue 1 protein. This sequence
XX CC is used to illustrate the method of the invention. (Updated on 23-OCT-
XX XX
XX SQ Sequence 266 AA;
XX
Query Match 97.3%; Score 107; DB 6; Length 266;
Best Local Similarity 95.2%; Pred. No. 4.9e-10;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
Db 122 AANARERRRMHGLNHAFDEL 142

Search completed: September 21, 2004, 21:20:57
Job time : 10.912 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:18:48 ; Search time 2.576 Seconds
(without alignments)
420.864 Million cell updates/sec

Title: US-09-980-381A-70
Perfect score: 110
Sequence: 1 AANARERRRMHGLNHAFDQLR 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	110	100.0	57	4	US-08-722-570-8
2	110	100.0	57	4	US-08-932-411A-8
3	84	76.4	57	4	US-08-722-570-9
4	84	76.4	57	4	US-08-932-411A-9
5	81	73.6	57	4	US-08-722-570-6
6	81	73.6	57	4	US-08-932-411A-6
7	81	73.6	352	1	US-08-552-142A-4
8	81	73.6	352	1	US-08-910-973-4
9	81	73.6	352	4	US-09-499-227-4
10	81	73.6	352	5	PCT-US95-05741-4
11	80	72.7	57	4	US-08-722-570-5
12	80	72.7	57	4	US-08-932-411A-5
13	80	72.7	156	1	US-08-552-142A-9
14	80	72.7	156	1	US-08-910-973-9
15	80	72.7	156	4	US-09-499-227-9
16	80	72.7	156	5	PCT-US95-05741-9
17	80	72.7	356	1	US-08-552-142A-15
18	80	72.7	356	1	US-08-910-973-15
19	80	72.7	356	3	US-09-234-332-7
20	80	72.7	356	3	US-09-234-332-8
21	80	72.7	356	4	US-09-499-227-15
22	80	72.7	357	1	US-08-552-142A-2
23	80	72.7	357	1	US-08-910-973-2
24	80	72.7	357	4	US-09-499-227-2
25	80	72.7	357	5	PCT-US95-05741-2
26	73	66.4	57	4	US-08-722-570-7
27	73	66.4	57	4	US-08-932-411A-7

28 73 66.4 379 1 US-08-552-142A-11 Sequence 11, Appl
29 73 66.4 380 1 US-08-552-142A-17 Sequence 17, Appl
30 73 66.4 381 1 US-08-910-973-11 Sequence 11, Appl
31 73 66.4 381 4 US-09-499-227-11 Sequence 11, Appl
32 73 66.4 382 1 US-08-910-973-17 Sequence 17, Appl
33 73 66.4 382 3 US-09-234-332-9 Sequence 9, Appl
34 73 66.4 382 4 US-09-499-227-17 Sequence 17, Appl
35 69 62.7 57 4 US-08-722-570-4 Sequence 4, Appl
36 69 62.7 57 4 US-08-932-411A-4 Sequence 4, Appl
37 69 62.7 214 4 US-08-722-570-2 Sequence 2, Appl
38 69 62.7 214 4 US-08-932-411A-2 Sequence 2, Appl
39 68 61.8 214 4 US-08-932-411A-20 Sequence 20, Appl
40 67 60.9 57 4 US-08-722-570-3 Sequence 3, Appl
41 67 60.9 57 4 US-08-932-411A-3 Sequence 3, Appl
42 67 60.9 103 1 US-08-552-142A-13 Sequence 13, Appl
43 67 60.9 237 1 US-08-910-973-13 Sequence 13, Appl
44 67 60.9 237 4 US-09-499-227-13 Sequence 13, Appl
45 67 60.9 244 1 US-08-910-973-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-722-570-8
; Sequence 8, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/722,570
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-722-570-8

Query Match 100.0%; Score 110; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21

Db 6 AANARERRRMHGLNHAFDQLR 26

```
RESULT 2
US-08-932-411A-8
; Sequence 8, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELEPHONE: (415) 781-1989
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
US-08-932-411A-8
Query Match 100.0%; Score 110; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 3
US-08-722-570-9
; Sequence 9, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
```

```
ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 536S
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
US-08-722-570-9
Query Match 76.4%; Score 84; DB 4; Length 57;
Best Local Similarity 81.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 4
US-08-932-411A-9
; Sequence 9, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
```

TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-9

Query Match 76.4%; Score 84; DB 4; Length 57;
Best Local Similarity 81.0%; Pred. No. 3.8e-07;
Matches 17; Conservative 1; Mismatches 3; Indels 0;

Qy 1 AANARERRMHGHNHAFDQLR 21
Db 6 AANARERRMHGHNHAFDQLR 26

RESULT 5
US-08-722-570-6

Sequence 6, Application US/08722570
Patent No. 6555337
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, QiuFu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,570
FILING DATE: 27-SEP-1996
CLASSIFICATION: 5365
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-722-570-6

Query Match 73.6%; Score 81; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0;

Qy 2 AANARERRMHGHNHAFDQLR 21
Db 7 AANARERRMHGHNHAFDQLR 26

RESULT 6

US-08-932-411A-6

Sequence 6, Application US/08932411A
Patent No. 6566496

GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, QiuFu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/932,411A
FILING DATE: 15-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/772,009
FILING DATE: 19-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/722,570
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-6

Query Match 73.6%; Score 81; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AANARERRMHGHNHAFDQLR 21
Db 7 AANARERRMHGHNHAFDQLR 26

RESULT 7

US-08-552-142A-4
Sequence 4, Application US/08552142A
Patent No. 5695995
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA

Query Match 73.6%; Score 81; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.2e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AANARERRMHGHNHAFDQLR 21
Db 7 AANARERRMHGHNHAFDQLR 26

ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-NOV-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FHCR-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-552-142A-4

Query Match 73.6%; Score 81; DB 1; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMGLNHAFLQRL 21
Db 107 ANARERNRMEGLNDALDSLRL 126

RESULT 8
US-08-910-973-4
Sequence 4, Application US/08910973
Patent No. 5795723
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/910,973
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532

FILING DATE: 30-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR-1-10958
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-973-4
Query Match 73.6%; Score 81; DB 1; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 ANARERRMGLNHAFLQRL 21
Db 107 ANARERNRMEGLNDALDSLRL 126
RESULT 9
US-09-499-227-4
Sequence 4, Application US/09499227
Patent No. 644463
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/09/499,227
FILING DATE: 05-August-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-09-499-227-4

Query Match      73.6%; Score 81; DB 4; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 ANARERRMHGLNHAFDQLR 21
      ||||| ||||| ||||| |||||
Db      107 ANARERNRMHGLNDALDSLR 126

RESULT 10
PCT-US95-05741-4
; Sequence 4, Application PC/TUS9505741
; GENERAL INFORMATION:
; APPLICANT: Weintraub, Harold
; APPLICANT: Lee, Jacqueline E.
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Hollenberg, Stanley M.
; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Gene
; TITLE OF INVENTION: and Protein
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05741
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Broderick, Thomas F.
; REGISTRATION NUMBER: 31,332
; REFERENCE/DOCKET NUMBER: FHCR-1-8504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-225-0709
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-05741-4

Query Match      73.6%; Score 81; DB 5; Length 352;
Best Local Similarity 80.0%; Pred. No. 9.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 ANARERRMHGLNHAFDQLR 21
      ||||| ||||| ||||| |||||
Db      107 ANARERNRMHGLNDALDSLR 126

RESULT 11
US-08-722-570-5
; Sequence 5, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
```

```
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 536S
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-722-570-5

Query Match      72.7%; Score 80; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 ANARERRMHGLNHAFDQLR 21
      ||||| ||||| ||||| |||||
Db      7 ANARERNRMHGLNAALDNLR 26

RESULT 12
US-08-932-411A-5
; Sequence 5, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
```

NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-932-411A-5

Query Match 72.7%; Score 80; DB 4; Length 57;
Best Local Similarity 80.0%; Pred. No. 1.8e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 7 ANARERNMHGHNAAALNLR 26

RESULT 13
US-08-552-142A-9
Sequence 9, Application US/08552142A
Patent No. 5695995
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (Neurod) Genes
TITLE OF INVENTION: and Proteins
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 02-NOV-1995
APPLICATION NUMBER: US/08/552,142A
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FHCR-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-552-142A-9

Query Match 72.7%; Score 80; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 106 ANARERNMHGHNAAALNLR 125

RESULT 14
US-08-910-973-9
Sequence 9, Application US/08910973
Patent No. 5795723
GENERAL INFORMATION:
APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/910,973
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR-1-10958
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-973-9

Query Match 72.7%; Score 80; DB 1; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGHNHAFDQLR 21
||||| ||||| ||||| ||||| |||||
Db 106 ANARERNMHGHNAAALNLR 125

RESULT 15
US-09-499-227-9
Sequence 9, Application US/09499227
Patent No. 644463
GENERAL INFORMATION:

APPLICANT: Tapscott, Stephen J.
APPLICANT: Olson, James M.
TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/499,227
FILING DATE: 05-August-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 156 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-499-227-9

Query Match 72.7%; Score 80; DB 4; Length 156;
Best Local Similarity 80.0%; Pred. No. 5.6e-06;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ANARERRMHGLNHAFDQLR 21
DB 106 ANARERRMHGLNHALDNLK 125

Search completed: September 21, 2004, 21:25:51
Job time : 2.576 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:24:21 ; Search time 9.576 Seconds

(without alignments)
704.240 Million cell updates/sec

Title: US-09-980-381A-70

Perfect score: 110

Sequence: 1 AANARERRMHGLNHAFDQLR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	110	100.0	21	16	US-10-373-249-5
2	110	100.0	57	8	US-08-722-570-8
3	110	100.0	57	12	US-10-425-259-8
4	110	100.0	161	13	US-10-004-717-60
5	110	100.0	259	13	US-10-004-717-66
6	110	100.0	351	13	US-10-004-717-11
7	110	100.0	351	13	US-10-004-717-46
8	110	100.0	351	16	US-10-373-249-1
9	110	100.0	351	16	US-10-373-249-2
10	110	100.0	354	13	US-10-004-717-2
11	110	100.0	354	13	US-10-004-717-58
12	110	100.0	354	16	US-10-373-249-3
13	110	100.0	354	16	US-10-373-249-4
14	107	97.3	266	13	US-10-004-717-40
15	98	89.1	46	13	US-10-004-717-62

16	91	82.7	45	13	US-10-004-717-54	Sequence 54, Appl
17	149	80.0	13	13	US-10-004-717-23	Sequence 23, Appl
18	88	80.0	151	13	US-10-004-717-36	Sequence 36, Appl
19	87	79.1	138	13	US-10-004-717-50	Sequence 50, Appl
20	87	79.1	138	13	US-10-004-717-52	Sequence 52, Appl
21	84	76.4	45	13	US-10-004-717-56	Sequence 56, Appl
22	84	76.4	57	8	US-08-722-570-9	Sequence 9, Appl
23	84	76.4	57	12	US-10-425-259-9	Sequence 9, Appl
24	84	76.4	312	13	US-10-004-717-64	Sequence 64, Appl
25	83	75.5	325	15	US-10-120-801-96	Sequence 96, Appl
26	82	74.5	189	13	US-10-004-717-17	Sequence 17, Appl
27	81	73.6	57	8	US-08-722-570-6	Sequence 6, Appl
28	81	73.6	57	12	US-10-425-259-6	Sequence 6, Appl
29	81	73.6	337	13	US-10-221-625-20	Sequence 20, Appl
30	81	73.6	337	13	US-10-004-717-9	Sequence 9, Appl
31	81	73.6	337	13	US-10-004-717-44	Sequence 44, Appl
32	81	73.6	352	16	US-10-654-102-6	Sequence 6, Appl
33	81	73.6	352	16	US-10-654-102-16	Sequence 16, Appl
34	81	73.6	352	16	US-10-654-102-27	Sequence 27, Appl
35	81	73.6	352	16	US-10-654-102-38	Sequence 38, Appl
36	81	73.6	352	16	US-10-654-102-43	Sequence 43, Appl
37	80	72.7	57	8	US-08-722-570-5	Sequence 5, Appl
38	80	72.7	57	12	US-10-425-259-5	Sequence 5, Appl
39	80	72.7	113	16	US-10-654-102-30	Sequence 30, Appl
40	80	72.7	113	16	US-10-654-102-41	Sequence 41, Appl
41	80	72.7	156	16	US-10-654-102-37	Sequence 37, Appl
42	80	72.7	216	16	US-10-654-102-23	Sequence 23, Appl
43	80	72.7	285	16	US-10-654-102-29	Sequence 29, Appl
44	80	72.7	330	13	US-10-004-717-33	Sequence 33, Appl
45	80	72.7	330	13	US-10-004-717-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1

US-10-373-249-5
; Sequence 5, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:
; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR
; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: drosophila
US-10-373-249-5

Query Match 100.0%; Score 110; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
DB 1 AANARERRMHGLNHAFDQLR 21

RESULT 2

US-08-722-570-8
; Sequence 8, Application US/08722570
; Publication No. US20030044887A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Oinfu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/722,570
APPLICATION NUMBER: US/08/722,570
FILING DATE: 27-SEP-1996
CLASSIFICATION: 5365
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: protein
MOLECULE TYPE: protein
US-08-722-570-8

Query Match 100.0%; Score 110; DB 8; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 3
US-10-425-259-8
Sequence 8, Application US/10425259
Publication No. US20030224431A1
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
Ma. Qifu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
STREET: Flehr Hobbach Test Albritton & Herbert LLP
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/425,259
FILING DATE: 29-Apr-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/932,411A
FILING DATE: 15-SEP-1997
APPLICATION NUMBER: US 08/772,009
FILING DATE: 19-DEC-1996
APPLICATION NUMBER: US 08/722,570
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:

NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: protein
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-425-259-8

Query Match 100.0%; Score 110; DB 12; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 6 AANARERRMHGLNHAFDQLR 26

RESULT 4
US-10-004-717-60
Sequence 60, Application US/10004717
Publication No. US20020192665A1
GENERAL INFORMATION:
APPLICANT: YANG, QI
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
FILE REFERENCE: P01899US4
CURRENT APPLICATION NUMBER: US/10/004,717
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/585,645
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: 60/176,993
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/137,060
PRIOR FILING DATE: 1999-06-01
NUMBER OF SEQ ID NOS: 69
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 60
LENGTH: 161
TYPE: PRT
ORGANISM: chicken
US-10-004-717-60

Query Match 100.0%; Score 110; DB 13; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRMHGLNHAFDQLR 21
Db 89 AANARERRMHGLNHAFDQLR 109

RESULT 5
US-10-004-717-66
Sequence 66, Application US/10004717
Publication No. US20020192665A1
GENERAL INFORMATION:
APPLICANT: YANG, QI
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
FILE REFERENCE: P01899US4

Best Local Similarity 100.0%; Fied. NO. 0e-03; Indels 0; Gaps 0;
Matches 21; Conservative Mismatches 0;

```

; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-373-249-4

Query Match      100.0%; Score 110; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 8.le-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 AANARERRRMHGLNHAFDQLR 21
        |||||||
Db      163 AANARERRRMHGLNHAFDQLR 183

RESULT 14
US-10-004-717-40
; Sequence 40, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGBBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DRAGNESS.
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01

```

Wed Sep 22 12:21:31 2004

; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Zebra Fish
US-10-004-717-40

Query Match 97.3%; Score 107; DB 13; Length 266;
Best Local Similarity 95.2%; Pred.No. 1.7e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
Db 122 AANARERRRMHGLNHAFDELR 142

RESULT 15
US-10-004-717-62
; Sequence 62, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Puffer Fish
US-10-004-717-62

Query Match 89.1%; Score 98; DB 13; Length 46;
Best Local Similarity 90.5%; Pred.No. 7e-08;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AANARERRRMHGLNHAFDQLR 21
Db 1 AANARERRRMHGLNKAFDELR 21

Search completed: September 21, 2004, 21:40:26
Job time : 9.576 secs

Result No.	Score	Query		ID	Description
		Match	Length		
1	1627.5	87.4	351	2	A56387
2	219.5	11.8	347	2	Q79999
3	214	11.5	352	2	I51687
4	213	11.4	312	2	A40708
5	210.5	11.3	381	2	G02668
6	207.5	11.1	381	2	A57059
7	205.5	11.0	337	2	I48682
8	205.5	11.0	337	2	I57038
9	202	10.8	357	2	Q4703
10	202	10.8	357	2	I49338
11	202	10.8	383	2	Q43688
12	199	10.7	192	2	T15764
13	199	10.7	381	2	Q43647
14	195	10.5	311	2	S20085
15	195	10.5	367	2	Q6087
16	184	9.9	71	2	T23978
17	179	9.6	331	2	A36358
18	178	9.6	329	2	A37864
19	175.5	9.4	267	2	A30988
20	174.5	9.4	326	2	S71755
21	174	9.3	278	2	A43814
22	174	9.3	278	2	S16678
23	171	9.2	334	2	G02409
24	167	9.0	207	2	I53154
25	162.5	8.7	245	2	B43814
26	157.5	8.5	198	2	A57717
27	153.5	8.2	237	2	T15548
28	152	8.2	147	2	T29241
29	148.5	8.0	170	2	T29971

RESULT 2
JC7999

basic helix-loop-helix transcription factor, Zath3 - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C:Date: 10-Nov-2003 #sequence_revision 10-Nov-2003 #text_change 24-Nov-2003
 C:Accession: JG7999
 R:Pack, S.H.; Yeo, S.Y.; Yoo, K.W.; Hong, S.K.; Lee, S.; Rhee, M.; Chitnis, A.B.; Kim, C.
 Biochem. Biophys. Res. Commun. 308, 184-190, 2003
 A:Title: Zath3, a neural basic helix-loop-helix gene, regulates early neurogenesis in the zebra fish.
 A:Reference number: JG7999; PMID:12890499
 A:Accession: JG7999
 A:Molecule type: mRNA
 A:Residues: 1-347 <PAR>
 A:Cross-references: GB:AF204240
 A:Experimental source: (Embryo)
 C:Comment: This protein, which belongs to the family of basic helix-loop-helix transcriptional ganglia, and the roles in early neurogenesis in zebra fish.
 C:Genetics:
 A:Gene: zath3
 C:Keywords: basic helix-loop-helix protein; neurogenesis

Query Match 11.8%; Score 219.5; DB 2; Length 347;
 Best Local Similarity 24.6%; Pred. No. 7.8e-08;
 Matches 81; Conservative 41; Mismatches 112; Indels 95; Gaps 12;

QY 64 WLAPTL---QGICTARAAQYLLHSPELGASAAAPRDEVDGRLVRRSSGGASSKSPG 120
 Db 20 WMDELLSQDGRTPETIGHYSLHRSNRGPLEIGS--EDMDEEEEBEEDDEEMGLDGEKAP- 76
 QY 121 PVKVRQLCKLKGWVVDLGCSSQRAPSSKQVNGVOKO---RRLAANARERRMHGLN 176
 Db 77 -----KRGPKKKKMTAKQEPFRARIKANAKERSMHGLN 113
 QY 177 HAPQLNRVPSFNNDKKLYETLQMAQIYINALSELLOTPSGGEQF----- 224
 Db 114 DALDNLRRVWPVCYSKTKLKIETLRLARNYIWLSEVLES---GQSPESHGFVEMLCCKG 170
 QY 225 -PPPPASCKSDHHILRTAAAYE-----GGAGNATAAGAAQASGSGSQRTPPGSCRTFSA 278
 Db 171 LPQTSNLVAGCQLQGLPTTMLKDEKGVGAGVPQGHPISYSPGLPSP----- 221
 QY 279 PASAGGYSVQLDALHGFSTFSDSALTAMMAQK-----NLSPS-----LPGSTLQPVQE 325
 Db 222 -----YCTMAASHLLHLKGYGPPYENSFNECSSGTPPYDGLPTPLPSI 266
 QY 326 ENS---KTSRPSHRSDGEFSPH-SHYSDS 350
 Db 267 SGNFALKQEPSPREAERNFTPHPTIYSS 295

neurogenic differentiation factor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: I51687
 R:Lee, J.E.; Hollenberg, S.M.; Snider, L.; Turner, D.L.; Lipnick, N.; Weintraub, H.
 Science 268, 836-844, 1995
 A:Title: Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein.
 A:Reference number: A56481; MUID:95273957; PMID:7754368
 A:Accession: I51687
 A>Status: preliminary;
 A:Molecule type: mRNA
 A:Residues: 1-352 <LEE>
 A:Cross-references: EMBL:U28067; NID:g854740; PIDN:AAC59675.1; PID:g854741
 C:Genetics:
 A:Gene: neuroD

Query Match 11.5%; Score 214; DB 2; Length 352;
 Best Local Similarity 29.5%; Pred. No. 1.9e-07;
 Matches 70; Conservative 31; Mismatches 86; Indels 50; Gaps 9;

QY 144 RQRPASSKQVNGVQ-----KQRLAANARERRMHGLNHAFLDQLRNVIPSFNNDKKLSKYE 199
 Db 83 KRRGPKKKKMTAKVERFVKRMKANARENRMHGLNDALDSLRLKVVPCYSKTKQKLSKIE 142

QY 200 TLQMAQIYINALSEL-----OTPSGGEQPPPP-----PASCK 232
 Db 143 TRLAKNYIWLSELRLSGKSPDLVSFVQTLCKLSQFTNLVAGCLQNLPRFLPEQSQ 202
 QY 233 SDHHLRTAASVEYGAGNATAAGAAQASGSGSQRTPP--GSCRTFRSAPASAGYSVQLDA 291
 Db 203 DIQSHMQTASS-----SFPLOGVPYQSPG--LSPPPYGTWDSHVFHKPHSVGAALPE 254
 QY 292 LHFSTFEDSALTAMMAQKNLSPSPGSIQPVQENSKTSPSRHS--DGEFSPHSY 347
 Db 255 F----FDSSTVT-----ECTSPSPDGLSPPLSVNGNFTPKHEHSEYDKNVTFTMHY 302

RESULT 4
 A40708
 basic-helix-loop-helix protein ato - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: A40708
 R:Jarman, A.P.; Grau, Y.; Jan, L.Y.; Jan, Y.N.
 Cell 73, 1307-1321, 1993
 A:Title: atonal is a proneural gene that directs chordotonal organ formation in the Drosophila embryo.
 A:Reference number: A40708; MUID:93313961; PMID:8324823
 A:Accession: A40708
 A>Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-312 <JAR>
 A:Cross-references: GB:L36646; NID:g551565; PIDN:AAA21879.1; PID:g551566
 A:Note: sequence extracted from NCBI backbone (NCBIN:135094, NCBI:P:135095)
 C:Genetics:
 A:Gene: FlyBase:ato
 A:Cross-references: FlyBase:FBgn0010433

Query Match 11.4%; Score 213; DB 2; Length 312;
 Best Local Similarity 34.0%; Pred. No. 1.9e-07;
 Matches 68; Conservative 21; Mismatches 85; Indels 26; Gaps 6;

QY 33 PPQPRA-----TLQAREHPYVPPLSLDSTDTP--RAWLAPTLOGICTARA 77
 Db 123 FVASPPAVEVMGSSNVGTCTIPASAAP--KPKRSYTKKNQPTTATSTPTAAESSASV 180
 QY 78 AQYLLHSPELGASAAAPRDEVDGRLVRRSSGGASSKSPGVKVRQLCKLKGWVV 137
 Db 181 NLYTEFFQNFDFNSALFDDSVDEDDLM-LFSGGDFDQNGSF-----DLADGENQ 232
 QY 138 DELGCSRQRPASSKQVNG-VQKORLLAANARERRMHGLNHAFLDQLRNVIPSFNNDKKLS 196
 Db 233 DAAAGSGKKRGKQITPVVKKRLLAANARERRMQNLAQFRLRQYLPCLGNDRQLS 292
 QY 197 KYETLQMAQIYINALSELIQ 216
 Db 293 KHETLQMAQTYISALGDLRL 312

RESULT 5
 G02668
 neurogenic basic-helix-loop-helix (bHLH) protein - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
 C:Accession: G02668
 R:Tapscott, S.J.; Tamimi, R.T.; McCormick, B.M.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: H01573
 A:Accession: G02668
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-381 <TAP>
 A:Cross-references: EMBL:U58681; NID:g1477748; PID:g1477749
 C:Genetics:
 A:Gene: NeuroD2

Query Match 11.3%; Score 210.5; DB 2; Length 381;

Query Match	11.0*	Score 205.5;	DB 2;	Length 337;
Best Local Similarity	29.0#;	Pred. No. 6.8e-07;		
Matches 71;	Conservative 31;	Mismatches 94;	Indels 49;	Gaps 7;

44 RGSIKAPAEETEKEEEDREEE-----DENGSLRRRLKKKTKTLRLERV 92
Db

158 KQRELAANARRRRMHGLNHFQDLNRVIPSFNNDKLSKYETLQMAQIVINALSELLQT 217
Qy

93 KFRFEANARERNRMHGLNDALNDLRKVVPCYSYTKLSKIETLRLLAKNIVWALSELRI 152
Db

218 PSQGEQP-----PPPPASCKSDHHHLRTAASYEGGAGNATACAGAAQASGGQ 264
Qy

153 ---GKRPDLTLFVQNLCKGLSQPTTNLVAGCLQL-NARSLFMGGQGEAAHHTRSPYSTFY 208
Db

265 RP-----TPGSCRTRFAPASAGYSVQLDHALHFSTFEDSDALTAMQAQKNLSPSLPG 317
Qy

209 PPYHSPELATPPGHGTLDNCKSMKPNYCYAYESFYESTGPECA-----SPOFEG 258
Db

Q1 259 PLSPP 263

Db

RESULT 8

I57038

gene Dlx-3 protein - mouse

C:Species: Mus sp. (mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997

C:Accession: I57038

R:Robinson, G.W.; Mahon, K.A.

Mech. Dev. 48, 199-215, 1994

A:Title: Differential and overlapping expression domains of Dlx-2 and Dlx-3 suggest dist

A:Reference number: I57038; MUID:95200802; PMID:7893603

A:Accession: I57038

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: mRNA

A:Residues: 1-337 <RES>

A:Cross-references: GB:S75299; NID:g896199; PID:g896200

C:Genetics:

A:Gene: Dlx-3

Query Match 11.0%; Score 205.5; DB 2; Length 337;

I49338
neurogenic differentiation factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence revision 16-Aug-1996 #text change 05-Nov-1999

```

Query Match      10.8%; Score 202; DB 2; Length 383;
Best Local Similarity 25.9%; Pred. NO. 1.4e-06;
Matches 98; Conservative 36; Mismatches 120; Indels 124; Gaps 17;

QY 18 DHHRQPPHLLPQPPPP-----QPPATLQAREHPVYPPELSLLDSTDPRAWLAP 67
      : ||||| : : : : :
Db 27 DEPRSDGADAPPQPPFAPGSGAPGPAARAKPVSLRGEE--IPE-----p 69
      : ||||| : : : : :

QY 68 TLQIGCTARAQAYLLHSPBGLGASAAAPRDEVDGRELVRSSGGASSKSPGPKVREQ 127
      : ||||| : : : : :

```

Db 70 TLAEV-----KEGELGGEEREEB-EGUDE-----AEGRPKKRGPKKRWMT 113
 QY 128 LCKLKGVVVDELGCSQRAPSSKQVNGVQKORLIANARERRMHGLNHAFDQLRNVIP 187
 Db 114 KARLER-----SKLRQKANARERNRMDLNAALDNLKRVVP 150
 QY 188 SFNNDKLSKYETLQMAQIYINALSELL-----OTPSGGEQPPPP--PASC-K 232
 Db 151 CYSKTQKLSKIETLRLAKNYIWLSEILRSKRPDLVSVQTLCKGLSQTNLVAGCLQ 210
 QY 233 SDHHHLRTAAAYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPPG 270
 Db 211 LNSRNFLEQADGAGRHGSGGPPAMHPYVPCSRLAGAACQAGGLGGGAHAALRTHG 270
 QY 271 SCRTFRGAPASAGGYSVOLDALHFTSFEDSALTAMMAQKNLSPSLPGSILQPVQENSKT 330
 Db 271 YCAAYETLYAAGGGAGSPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPD 319
 QY 331 SPRSHRSDGEFSPSHYS 348
 Db 320 HEKSY-----HYS 327

RESULT 12

TL5764

hypochemical protein C34E10.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: TL5764

R:Kirsten, J.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of C. elegans cosmid C34E10.

A:Reference number: Z18399

A:Accession: TL5764

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-192 <KIR>

A:Cross-references: EMBL:U10402; NID:g500723; PID:g500730; PIDN:AAA19069.1; CESP:C34E10.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C34E10.7

A:Introns: 65/2; 107/3

Query Match

Best Local Similarity 10.7%; Score 200; DB 2; Length 192;

Matches 68; Conservative 13; Mismatches 76; Indels 44; Gaps 6;

QY 157 OKORLIANARERRMHGLNHAFDQLRNVIPSFNNDKLSKYETLQMAQIYINALSELLQ 216

Db 17 RKVRVKANGREARMHGLNMLDMREVIPITTOHKLKSKLETLEARNYIDALQRMLO 76

QY 217 TPSGGEQPPPPASCKSDHHHLRTAASVEGGAGNATAAGAAQASGGSQRTPPGSCRTRF 276

Db 77 T---NEQPTP-----LEYAHTLANGLSQTTNNLANLLQVQRLPLPPSQFOIF 122

QY 277 SAPASAGGYSVOLDALH-----FTFEDSALTAMMAQKNLSPSLPGSILQPVQENSKTSPR 333

Db 123 SDPSHH-----QLHPSHPHPHSSF-----SSSPSSSCSPPPYYGPTQPS 163

QY 334 SHRSDEGFS-----HSH 346

Db 164 AAPLOGSCDPQYQMYHQHSH 184

RESULT 13

JC4647

KWS protein - rat

N:Alternate names: basic helix-loop-helix protein homolog

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000

C:Accession: JC4647

R:Kume, H.; Maruyama, K.; Tomita, T.; Iwatsubo, T.; Saido, T.C.; Obata, K.

Biochem. Biophys. Res. Commun. 219, 526-530, 1996

A:Title: Molecular cloning of a novel basic helix-loop-helix protein from the rat brain.
 A:Reference number: JC4647; MUID:96193685; PMID:8605021
 A:Accession: JC4647
 A:Molecule type: DNA
 A:Residues: 1-381 <KDM>
 A:Cross-references: DDBJ:D82868; NID:g1166397; PIDN:BAAL1615.1; PID:g1166398
 A:Experimental source: brain
 C:Comment: This protein is involved in synaptic plasticity, and has a role specific to n-loop-helix domain.
 C:Keywords: brain
 F:122-134/region: basic

Query Match 10.7%; Score 199; DB 2; Length 381;

Best Local Similarity 26.1%; Pred. No. 2.2e-06;

Matches 81; Conservative 31; Mismatches 86; Indels 112; Gaps 12;

QY 18 DHHQPOPHHLPQPPPP-----OPPATLOAREHPVVPPELSLLDSTDRAWLAP 67

Db 27 DEPRSDKGADAPPQPPPPAPFGSGAPPARATKCVSLRGEVP-----EP 68

QY 68 TLQGICTARAAQYLLHSPELGASAAAPRDVDGREGELVRRSSGGASSSKSPGVKREQ 127

Db 69 TLAEV-----KEGELGGEEREEB-EGUDE-----AEGRPKKRGPKKRWMT 112

QY 128 LCKLKGVVVDELGCSQRAPSSKQVNGVQKORLIANARERRMHGLNHAFDQLRNVIP 187

Db 113 KARLER-----SKLRQKANARERNRMDLNAALDNLKRVVP 149

QY 188 SFNNDKLSKYETLQMAQIYINALSELLQTSFGSGEQPPPPASCKSDHHHLRTAASVEGG 247

Db 150 CYSKTQKLSKIETLRLAKNYIWLSEILRS---GKRP-----DLVSY--- 188

QY 248 AGNATAAGAAQASGGSQRTPPGSCRTFRSAPASAGYSVOLDALHFTFED- -SALT 303

Db 189 -----VQLCKLGLSQPT-----TNLVAGC-----LQNSRNFLEQADGAAFT 228

QY 304 AMMAQKNLSP 313

Db 229 ARVARSAMHP 238

RESULT 14

S20085

transcription factor SCL - chicken

C:Species: Gallus gallus (chicken)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999

C:Accession: S20085

R:Goodwin, G.; MacGregor, A.; Zhu, J.; Crompton, M.R.

Nucleic Acids Res. 20, 368, 1992

A:Title: Molecular cloning of the chicken SCL cDNA.

A:Reference number: S20085; MUID:92158629; PMID:1741264

A:Accession: S20085

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-311 <GOO>

A:Cross-references: EMBL:X63371; NID:g62844; PIDN:CAA44971.1; PID:g62845

C:Superfamily: lyl-1 protein

C:Keywords: DNA binding; transcription regulation

Query Match 10.5%; Score 195; DB 2; Length 311;

Best Local Similarity 24.6%; Pred. No. 3.2e-06;

Matches 77; Conservative 37; Mismatches 121; Indels 76; Gaps 11;

QY 29 POPPPPPQPPATLOAREHPVVPPELSLLDSTD-PRAWLAPTLOGICTARAAQYLLHSPEL 87

Db 7 PAPPPSSDPR--DARRH---DPEADATSEPDSSRGGMEFPAPQLLLNGAAKEAGRPS 61

QY 88 GASEAAAPDEVDGGRG---ELVRRSSGGASSSKSPG--PVKVRQQLCKLK----- 132

Db 62 GPAAAVPVIEIIVRRGSLDIKSREAAGEAMORAPGAEPCAAEEACAAECARVQLSPALP 121

QY 133 ----GGVVVDELG-----CSRQAPSSKQVNG---VOKQR 161

Db 122 LQPPGRAMLYNLGQPLTIGSGFTGEPDSFMSYGNRVKRRPSPYEMEITDGPHTKVVR 181

QY 162 LAANARERRRMHGLNHFADOLRNVIPSFNNDKLSKYETLQMAQIYINALSELL--OTPS 219

Db 182 IFNSRRERWQNVNGFAELRLKLIPTHPDPKLSKNEILRLAMKIINFLAKLNDQEE 241

QY 220 GGEOPPPPPACKSDHHHLRTAASYEGGAGNATAAGAAQAGSGSQRPPTPGSCRTFSAP 279

Db 242 GNQR-----GKVNKD-SGIVQEDLLQDMLSPNSSCGSSLDGA 277

QY 280 ASAGGYSVQLDAL 292

Db 278 ASDSFTEHDTL 290

RESULT 15

JC6087

helix-loop-helix transcription factor, BETA3 - hamster

C:Species: Cricetinae Gen. sp. (hamster)

C>Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999

C:Accession: JC6087

R: Peyton, M.; Stellrecht, C.M.M.; Naya, F.J.; Huang, H.P.; Samora, P.J.; Tsai, M.J.

Mol. Cell. Biol. 16, 626-633, 1996

A:Title: BETA3, a novel helix-loop-helix protein, can act as a negative regulator of BETA3

A:Reference number: JC6087; MUID:96140430; PMID:8552091

A:Accession: JC6087

A:Molecule type: mRNA

A:Residues: 1-367 <PEY>

A:Cross-references: GB:580870; NID:G1911496; PIDN:AAB50691.1; PID:G1911497

A:Experimental source: insulin tumor cell

C:Comment: This factor belongs to the tissue-specific class B basic helix-loop-helix family.

C:Genetics:

A:Gene: beta3

C:Keywords: repressor; transcription factor

Query Match 10.5%; Score 195; DB 2; Length 367;

Best Local Similarity 32.8%; Pred. No. 3.9e-06;

Matches 79; Conservative 18; Mismatches 84; Indels 60; Gaps 10;

QY 67 PTLOGICTARAAQYLLHSPBLGASEA-----APDEVDRGELVRRSSGGASSSKSP 119

Db 116 FNLSSL-PAGAAALCKYGESAGRGSAVSSGGEGSQSPDDSDGRCELVR-AGGADPRASP 173

QY 120 GPVKVREQLCKLXGVVVDLGGCSQR-----APSSKQVN 154

Db 174 GAG-----GGTKVVE-GCSNAHLHGAGLPPCGSTGSGGGGGGGGGSSSKKS 223

QY 155 GVQKORRLAANARERRRMHGLNHFADOLRNVIPSFNND--KLSKYETLQMAQIYI----- 208

Db 224 KEQKALRLNINARERRRMHGLNHFADOLRNVIPSFNND--KLSKYETLQMAQIYI----- 283

QY 209 NALSELL-----QTPSGGQPPPPPPASCKSDHHHLRTAASYEGGAGNATAAGAAQQA 259

Db 284 QALEERRLVAYLNQQAISASLPSSAAAAAALHPALGA-YEQAGYPFSAGLPPA 342

QY 260 S 260

Db 343 A 343

Search completed: September 21, 2004, 21:25:00

Job time : 50.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:35:26 ; Search time 27.376 Seconds
(without alignments)
673.321 Million cell updates/sec

Title: US-09-980-381A-58

Perfect score: 1862

Sequence: 1 MSRLHAEAEVKEVLGDHH.....HRSDGEFSPHSHYSDEAS 354

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1862	100.0	354	1 ATH1 HUMAN	Q92858 homo sapien
2	1627.5	87.4	351	1 ATH1 MOUSE	P48985 mus musculus
3	228.5	12.3	198	1 AMOS DROME	Q9Y0A7 drosophila
4	222	11.9	250	1 NGN2 HUMAN	Q9H2A3 homo sapien
5	219	11.8	208	1 NGN1 BRAPE	O42506 brachydanio
6	217	11.7	263	1 NGN2 MOUSE	P70447 mus musculus
7	214	11.5	352	1 NDF1 XENLA	Q91616 xenopus lae
8	213	11.4	312	1 ATO DROME	P48987 drosophila
9	210	11.3	382	1 NDF2 HUMAN	Q15784 homo sapien
10	209.5	11.3	237	1 NGN1 HUMAN	Q92886 homo sapien
11	207.5	11.1	355	1 NDF1 MESAU	O60430 mesocricetu
12	205.5	11.0	337	1 NDF6 MOUSE	P48986 mus musculus
13	205	11.0	244	1 NGN1 RAT	P70595 rattus norv
14	204	11.0	382	1 NDF2 RAT	Q63689 rattus norv
15	202	10.8	244	1 NGN1 MOUSE	P70660 mus musculus
16	202	10.8	357	1 NDF1 MOUSE	Q60867 mus musculus
17	202	10.8	357	1 NDF1 RAT	Q64289 rattus norv
18	202	10.8	383	1 NDF2 MOUSE	Q62414 mus musculus
19	201.5	10.8	214	1 NGN3 MOUSE	P70661 mus musculus
20	200.5	10.8	337	1 NDF6 HUMAN	Q961K8 homo sapien
21	200	10.7	192	1 YLB7 CAEEL	P46581 caenorhabdi
22	199	10.7	356	1 NDF1 HUMAN	Q13562 homo sapien
23	197.5	10.6	357	1 NDF1 CHICK	P79765 gallus gall
24	195	10.5	311	1 SCL CHICK	P24899 gallus gall
25	195	10.5	367	1 BET3 MESAU	O09029 mesocricetu
26	192.5	10.3	214	1 NGN3 HUMAN	Q9Y4Z2 homo sapien
27	191.5	10.3	330	1 NDFM CHICK	P79766 gallus gall
28	191.5	10.3	331	1 NDF4 HUMAN	Q9Hd90 homo sapien
29	187	10.0	330	1 NDF4 MOUSE	O09105 mus musculus
30	184	9.9	71	1 L132 CAEEL	Q10574 caenorhabdi
31	183	9.8	298	1 OLG2 CHICK	Q902b3 gallus gall
32	179	9.6	331	1 TAL HUMAN	P17542 homo sapien
33	178	9.6	329	1 TAL_MOUSE	P22091 mus musculus

RESULT 1

ID	ATH1 HUMAN	STANDARD;	PRT;	354 AA.
AC	Q92858;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Atonal protein homolog 1 (Helix-loop-helix protein hATH-1).			
GN	ATOHL OR ATH1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=97026280; PubMed=8872459;			
RA	Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,			
RA	Zoghbi H.Y.;			
RT	"Evolutionary conservation of sequence and expression of the bHLH			
RT	protein Atonal suggests a conserved role in neurogenesis.";			
RL	Hum. Mol. Genet. 5:1207-1216(1996).			
CC	!- FUNCTION: Activates E box-dependent transcription in collaboration			
CC	with E47, but the activity is completely antagonized by the			
CC	negative regulator of neurogenesis HES1. May play a role in the			
CC	differentiation of subsets of neural cells by activating E box-			
CC	dependent transcription (By similarity).			
CC	!- SUBUNIT: Efficient DNA binding requires dimerization with another			
CC	bHLH protein.			
CC	!- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U61148; AABA1305.1; -			
DR	TRANSFAC; T04544; -			
DR	Genew; HGNC:797; ATOHL.			
DR	MIM; 601461; -			
DR	GO; GO:0003700; P:transcription factor activity; TAS.			
DR	GO; GO:0007417; P:central nervous system development; TAS.			
DR	GO; GO:0006366; P:transcription from Pol II promoter; TAS.			
DR	InterPro; IPR001092; HLH_basic.			
DR	Pfam; PF00010; HLH; 1.			
DR	SMART; SM00353; HLH; 1.			
DR	PROSITE; PS00888; HLH; 1.			
KW	Transcription regulation; Activator; DNA-binding; Nuclear protein.			
FT	DOMAIN 29 38 POLY-PRO.			
FT	DNA_BIND 160 171 BASIC DOMAIN.			
FT	DOMAIN 172 212 HELIX-LOOP-HELIX MOTIF.			
FT	DOMAIN 224 228 POLY-PRO.			
SQ	SEQUENCE 354 AA; 38160 MW; AB12F1E917A00A8D CRC64;			

ALIGNMENTS

34	177.5	9.5	267	1	LYL1_HUMAN	P12980 homo sapien
35	177	9.5	315	1	NDF4_XENLA	P79920 xenopus lae
36	174	9.3	278	1	LYL1_MOUSE	P27792 mus musculus
37	171	9.2	323	1	OLG2_HUMAN	Q13516 homo sapien
38	170	9.1	263	1	ASC2_MOUSE	O35885 mus musculus
39	170	9.1	398	1	TAP_DROME	O16867 drosophila
40	169	9.1	323	1	OLG2_MOUSE	Q9eqw6 mus musculus
41	167	9.0	207	1	SCX_MOUSE	Q64124 mus musculus
42	165	8.9	199	1	TF15_HUMAN	Q12870 homo sapien
43	164	8.8	195	1	TF15_MESAU	Q60539 mesocricetu
44	158	8.5	195	1	TF15_MOUSE	Q60756 mus musculus
45	154.5	8.3	206	1	MUSC_HUMAN	O60682 homo sapien

Db 298 ALTANMAQKOLSPSLPGILQVQEDNSKTSRSHRSDGFSPHSHYSDSEAS 351

RESULT 3

ID AMOS DROME STANDARD; PRT; 198 AA.

AC Q9Y0A7; Q9VJ76;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Basic helix-loop-helix transcription factor Amos (Reduced olfactory organs protein) (Rough eye protein) (Absent MD neurons and olfactory sensilla protein) (Amos protein).

DE GN AMOS OR ROI OR ROLO OR CG10393.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A., AND FUNCTION.

RC STRAIN=Oregon-R;

RX MEDLINE=20170246; PubMed=10707973;

RA "Amos, a proneural gene for Drosophila olfactory sense organs that is regulated by lozenge."

RT Neuron 25:69-78(2000).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amaratunga P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berhan P.P., Bhandari D., Bolshakov S., Botchan M.R., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de la Fabre B., Delecher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster K.G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

RN [3]

RP FUNCTION.

RX MEDLINE=20170245; PubMed=10707972;

RA Huang M.L., Hsu C.H., Chien C.T.;

"The proneural gene amos promotes multiple dendritic neuron formation in the Drosophila peripheral nervous system.";

RT Neuron 25:57-67(2000).

RL

CC -!- FUNCTION: Transcription factor involved in early neurogenesis.

CC Promotes multiple dendritic (MD) neuron formation. Required for olfactory sensilla.

CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another bHLH protein. Interacts with Daughterless.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- DEVELOPMENTAL STAGE: During embryonic development, amos is expressed in patches of ectodermal cells, and the expression is quickly restricted to sensory organ precursors.

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

CC

DR EMBL; AF166113; AAD45410.1; -.

DR EMBL; AE003659; AAF53678.1; -.

DR HSSP; P25912; 1HLO.

DR FlyBase; FBgn0003270; amos.

DR InterPro; IPR001092; HLH_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS00888; HLH; 1.

DR Nuclear protein; Transcription regulation; Developmental protein;

KW Neurogenesis.

FT DOMAIN 111 128 SER-RICH

FT DOMAIN 138 150 BASIC DOMAIN

FT DOMAIN 151 191 HELIX-LOOP-HELIX MOTIF.

FT CONFLICT 25 25 S -> G (IN REF. 1).

SQ SEQUENCE 198 AA; 22569 MW; F1D22AF2BA1C670B CRC64;

Query Match 12.3%; Score 228.5; DB 1; Length 198;

Best Local Similarity 34.6%; Pred. No. 6.5e-08;

Matches 73; Conservative 22; Mismatches 57; Indels 59; Gaps 7;

QY 34 PPQPPA-----TLQAREHPVPPPELSLDSTDPRA-----WLAPT---LOGICT 74

Db 14 PDEAPAIPEFLSNDTFQLEQLMYQEFSTDSQSGANSCELEMYDTPSVLEHMLN 73

QY 75 AR-AAQYLLHSPFELGASAAAPR-----DEVDCRGELVRRSSGGASSKSPGVKVR 125

Db 74 AQEQOOHHLQANPLGNQGRSPRYWNKQGRSKFYDKLSTMSSTSSSSSS----- 126

QY 126 EQLCXKLGKGVVDELGCSQRAPSSKQVNG-VOKORRLAANARERRRMHGLNHPDQLRN 184

Db 127 -----SSAGTGGVLEKGRRLAANARERRRMHGLNHPDQLRN 163

QY 185 VIPSFNNDKLSKYETLQMAQIYINALSELL 215

Db 164 VVPSLGHDRRLSKYETLQMAQYIGDLVTLL 194

RESULT 4

NGN2 HUMAN STANDARD; PRT; 250 AA.

ID NGN2 HUMAN

AC Q9H2A3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Neurogenin 2 (Fragment).

GN NEUROG2 OR NGN2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

```
RP SEQUENCE FROM N.A.
RA Simmons A.D., Horton S., Abney A.L., Johnson J.E.;
RT "Neurogenin 2 expression in ventral and dorsal spinal neural tube
RL progenitor cells is regulated by distinct enhancers.";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in neurogenesis (By similarity).
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF303002; AAG40770.1; -.
CC Genew; HGNC:13805; NEUROG2.
CC MIM; 606624; -.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC DR PROSITE; PS50888; HLH; 1.
CC DR Neurogenesis; Developmental protein, Differentiation.
CC KW NON_TER 1
CC FT DNA_BIND 100 111 BASIC DOMAIN.
CC FT DOMAIN 112 152 HELIX-LOOP-HELIX MOTIF.
CC SQ SEQUENCE 250 AA; 26130 MW; BB381FF5EF1500EF CRC64;
CC
CC Query Match 11.9%; Score 222; DB 1; Length 250;
CC Best Local Similarity 30.4%; Pred. No. 2.1e-07;
CC Matches 90; Conservative 27; Mismatches 89; Indels 90; Gaps 14;
CC
QY 84 SPELGA-----SAAAPRDEVDGRLVRRSSGGASSKSPGVKVRQELCKLG--GVV 136
Db 2 SPALAAITPLSSADEEEEPGASGARRQRAEAGCGAGGVAAGAECEPARLLGLV 61
QY 137 VBLEGC-----SKQNGVKQRRLAANAFERRRMHGLHAFQDLRNVIP 187
Db 62 HD---CKRPPSRARVSRGAKTAETVQRKTKRLKANNRRNRNMENLNAALDALREVL 118
QY 188 SFNNDKLSKYETLQMAQIYINALSELQTPSGEGQPPPPASCKSDHHHLRTAAVEGG 247
Db 119 TTPEDAKTKIETLRFHNYIINALTETLR-----LADHCGG 154
QY 248 AGNAT-----AAGAQAQSGGSGRFTPPGSCRTFRSAPAGGYGVQLDAL 292
Db 155 GGGGLPGALFSEAVLLSPGASAAALSSGDSFSPASTWSC-TNSPAPSSS----- 203
QY 293 HSTFEDSALTAMAAQKNLSPSLP-GSIL---QPQVEENSKYSPRHSRSDGRFSPH 344
Db 204 -VSSNSTSPSYC-----TLSPASPGSDMDYQWP-----PPDQKHR-----YAPH 242
CC
RESULT 5
NGN1_BRARE
ID NGN1_BRARE STANDARD; PRT; 208 AA.
AC O42606;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenin related protein-1).
GN NEUROG1 OR NEUROD3 OR NGN1 OR NGRI.
OS Brachydanio rerio (Zebrafish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
```

```
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RX MEDLINE=98072331; PubMed=9409673;
RA Blader P., Fischer N., Gradwohl G., Guillemont F., Straetle U.;
RT "The activity of neurogenin1 is controlled by local cues in the
RL zebrafish embryo.";
RL Development 124:4557-4569 (1997).
CC [2]
CC SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
CC STAGE.
CC TISSUE=Embryo;
CC MEDLINE=98129571; PubMed=9469669;
CC Kim C.-H., Bae Y.-K., Yamanaka Y., Yamashita S., Shimizu T., Fujii R.,
CC Park H.-C., Yeo S.-Y., Huh T.-L., Hibi M., Hirano T.;
CC "Overexpression of neurogenin induces ectopic expression of HuC in
RT zebrafish.";
RL Neurosci. Lett. 239:113-116 (1997).
CC [3]
CC SEQUENCE FROM N.A.
CC Thisse B.;
CC "Danio rerio neurogenin the earliest marker of primary neurons starts
RT to be expressed at midgastrula.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC [4]
CC SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
CC TISSUE=Embryo;
CC MEDLINE=98401940; PubMed=9733104;
CC Korzh V., Sleptsova I., Liao J., He J., Gong Z.;
CC "Expression of zebrafish bHLH genes ngn1 and nrd defines distinct
RT stages of neural differentiation.";
RL Dev. Dyn. 213:92-104 (1998).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein (By similarity).
CC -1- TISSUE SPECIFICITY: Embryonic nervous system.
CC -1- DEVELOPMENTAL STAGE: First detected at the late gastrulation stage
CC (9 hours). Strongly expressed in distinct domains in the neural
CC plate at the 3-somite stage. At 24 hours, expressed in specific
CC Regions of the developing brain and in the spinal cord.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF024535; AAB82271.1; -.
CC EMBL; AF017301; AAB70528.1; -.
CC EMBL; AF036149; AAB88821.1; -.
CC EMBL; U94588; AAD09254.1; -.
CC ZFIN; ZDB-GENE-990415-174; neurogl.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS50888; HLH; 1.
CC DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DNA_BIND 71 82 BASIC DOMAIN.
FT DOMAIN 83 123 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 208 AA; 22911 MW; CDB333AE483DC51A CRC64;
CC
CC Query Match 11.8%; Score 219; DB 1; Length 208;
CC Best Local Similarity 30.1%; Pred. No. 2.7e-07;
CC Matches 69; Conservative 32; Mismatches 76; Indels 52; Gaps 8;
CC
QY 97 DEVDGRGELVRRSSGGASSKSPGVKVRQELCKLGKGVGVVDELGCSQRAPSSKQNGV 156
Db 21 DDEDSRSSL-HPASPASGCKPPASGLQKKRRRG-----RARNETTTHVV 67
```


QY 157 QKORLAARERRRMHGHAFDOLRVNIPSFNNKKLSKYETLQMAQIYINALSELLO 216
 DB 68 KKNRLKANDRRNRHMLNDALDRLSVLPFPDDTKLTKTETLRFHNYIWLSETIR 127
 QY 217 ---TPSGGEQPPPP---PPASCKSDHHHLRTAASVEGGAGNATAAGAQA---SGSGSRPT 267
 DB 128 IADQKQKSRDCPLLLPLGLSCWAD-----APSPGSDSCSWPFGASSSS 171
 QY 268 PPGSCRTFSPASAGGYS-VOLDALHFSTFEDSALTAMMAQKNLSPL 315
 DB 172 SPSYCNDSPPGSPAAMDFFGLQTDVY-----SYRNFVPSI 207

RESULT 6
 NGN2 MOUSE
 ID NGN2 MOUSE STANDARD; PRT; 263 AA.
 AC P70447; P70237;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenin 2 (Atonal protein homolog 4) (Helix-loop-helix protein
 DE MATH-4A) (MATH4A).
 GN NEUROG2 OR NGN2 OR ATOH4 OR ATH4A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97153565; PubMed=9000438;
 RA Sommer L., Ma Q., Anderson D.J.,
 RT "Neurogenins, a novel family of atonal-related bHLH transcription
 RT factors, are putative mammalian neuronal determination genes that
 RT reveal progenitor cell heterogeneity in the developing CNS and PNS."
 RL Mol. Cell. Neurosci. 8:221-241(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1;
 RX MEDLINE=971105839; PubMed=8948587;
 RA Gradowohl G., Fode C., Guillemot F.,
 RT "Restricted expression of a novel murine atonal-related bHLH protein
 RT in undifferentiated neural precursors."
 RL Dev. Biol. 180:227-241(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Simmons A.D., Horton S., Abney A.L., Johnson J.E.,
 RT "Neurogenin 2 expression in ventral and dorsal spinal neural tube
 RT progenitor cells is regulated by distinct enhancers."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U76207; AAC53028.1; -
 DR EMBL; Y07621; CAAG8900.1; -
 DR EMBL; AF303001; AAC40769.1; -
 DR MGD; MGI:109619; Neurog2.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.
 DR SMART; SM00353; HLH; 1.
 DR PROSITE; P55088; HLH; 1.

KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
 KW Neurogenesis; Developmental protein; Differentiation.
 FT DNA_BIND 113 124 BASIC DOMAIN.
 FT DOMAIN 125 165 HELIX-LOOP-HELIX MOTIF.
 FT CONFLICT 60 60 E -> G (IN REF. 2).
 SQ SEQUENCE 263 AA; 28215 MW; 817EF8246BD8CABE CRC64;

Query Match 11.7%; Score 217; DB 1; Length 263;
 Best Local Similarity 28.7%; Pred. No. 4.6e-07;
 Matches 98; Conservative 25; Mismatches 89; Indels 130; Gaps 15;

QY 40 TLQAREHFVYPPELSLLDSTDTP-RAWLAP-----TLQGITAPAAQYL 81
 DB 7 TLELKEE---BEVLMILGASAPASATLTPMSSADEEDEELRRPQSGARGQAEQGV 63
 QY 82 LHSPELGASEAAAAPRDEVDGRLVRRSGGASSSKSPGVKVRQLCKLKGVVVDELG 141
 DB 64 QGSPASGA-----GGC-----RFGRLGLMHECKRRP----- 90
 QY 142 CSRQRAPS-----SKOVNGVQKORBLAANARERRRMHGLNHAFDOLRVNIPFNNDKKLS 196
 DB 91 -SRSAVRGAKTAEIVQRIKTKRNNRRNMHLNAAALDALREVLPFPEDAKLT 149
 QY 197 KYETLQMAQIYINALSELLOTPSGGEGQPPPPSPASCKSDHHHLRTAASVEGGAGNATAA-- 254
 DB 150 KIETLRFHNYIWLSETIR-----LADHCAGAGGLQALF 185
 QY 255 -----GAQQASGGSGRPTTPGSCRTFSPASAGYSVQLDALHFSTFEDSALTAM 305
 DB 186 TEAVLLSPGAALGASGSPSPSSWSCT---NSPASSN-----STSPYSC----- 228
 QY 306 MAQKNLSPLSGSIL---QPVOEENSKTSPRSHRSDGFSFPH 344
 DB 229 ----TLSPASGSDVDYQWQ-----PPPEKHR-----YAPH 255

RESULT 7
 NDPL_XENLA
 ID NDPL_XENLA STANDARD; PRT; 352 AA.
 AC Q91616;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic differentiation factor 1 (NeuroD1).
 GN NEUROD1 OR NEUROD.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95273957; PubMed=7754368;
 RA Lee J.E., Hollenberg S.M., Snider L., Turner D.L., Lipnick N.,
 RA Weintraub H.,
 RT "Conversion of Xenopus ectoderm into neurons by NeuroD, a basic
 RT helix-loop-helix protein."
 RL Science 268:836-844(1995).
 CC -!- FUNCTION: Acts as a differentiation factor during neurogenesis.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in differentiating neurons of
 CC both the central and peripheral nervous systems.
 CC -!- DEVELOPMENTAL STAGE: Expressed during embryonic development.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28067; AAC59675.1; -.
DR PIR; I51687; I51687.
DR InterPro; IPR01092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
DR DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DOMAIN 59 78 GLU-RICH (ACIDIC).
FT DOMAIN 88 94 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 103 114 BASIC DOMAIN.
FT DOMAIN 115 155 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 352 AA; 39662 MW; 226298DB348233E CRC64;

Query Match 11.58; Score 214; DB 1; Length 352;
Best Local Similarity 29.5%; Pred. No. 9.9e-07;
Matches 70; Conservativity 31; Mismatches 86; Indels 50; Gaps 9;

QY 144 RORAPSSKQVNGVQ---KORLLAANARRRRMHGLNHLAFDQLNRNVPSPNNKKLSKYE 199
DB 83 KRGRPKKKKTKARVERFKYRMKANARRNRMHGLNDALDSLRLKVVPCYSKTQKLSKIE 142
QY 200 TLQWAIYINWALSELL-----QTPSGEQPPPP-----PASCK 232
DB 143 TLRLAKNYINWALSILRSKSPDLVSFVQTLCKLSQPTTNLVAGCLQLNPRTPFLPQSQ 202
QY 233 SDHHHLRTAASYEGGAGNATAAGQAQAGSGSQPTTPP-GSCRTFFSPAPAGGYSVQVDA 291
DB 203 DIQSHMGTASS-----SFLQGYFYQSPG--LPSPPYGTMDSHVPHKPSVGAALP 254
QY 292 LHFSTFSDSALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRS-DGEFFSPHSY 347
DB 255 F-----FDSSIVT-----ECTSPSPDGLSPFLSVNGNFTFKHSEYDKNYTTFMHY 302

RESULT 8
ATO DROME STANDARD; PRT; 312 AA.
AC P48987; Q9VHU0.
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Atonal protein.
GN ATO OR CG7508.
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Oregon-R;
RC MEDLINE=93313961; PubMed=9324823;
RA Jarman A.P., Grau Y., Jan L.Y., Jan Y.N.;
RT "Atonal is a proneural gene that directs chordotonal organ formation
in the Drosophila peripheral nervous system.";
RL Cell 73:1307-1321(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Stutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J.J., Brokstein P., Brottier P.,

```

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali B., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Klamel M.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RN FUNCTION.
RP MEDLINE=94255014; PubMed=8196767;
RX Jarman A.P., Grell E.H., Ackerman L., Jan L.Y., Jan Y.N.;
RA "Atonal is the proneural gene for Drosophila photoreceptors.";
RT Nature 369:398-400(1994).
CC -!- FUNCTION: Developmental protein involved in neurogenesis. Required
CC for the formation of chordotonal organs and photoreceptors. Seems
CC to bind to E boxes. Specifically required for the photoreceptor R8
CC selection.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Forms a heterodimer with Daughterless.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Proneural clusters and sense organ precursors
CC of the chordotonal organs, optic trow of the eye-antennal disk
CC and developing brain lobe.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L36646; AAA21879.1; -.
DR EMBL; AE003678; AAF54209.1; -.
DR PIR; A40708; A40708.
DR FlyBase; FBgn0010433; ato.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:0007173; P:EGF receptor signaling pathway; IGI.
DR GO; GO:0007605; P:hearing; IMP.
DR GO; GO:0007438; P:neocortex development; IMP.
DR GO; GO:0007422; P:peripheral nervous system development; NAS.
DR GO; GO:0045464; P:R8 cell fate specification; NAS.
DR GO; GO:0007224; P:smoothed signaling pathway; IGI.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Neurogenesis; Differentiation; Developmental protein; Nuclear protein;
KW Transcription regulation; DNA-binding
FT DNA_BIND 255 267 BASIC DOMAIN
FT DOMAIN 268 308 HELIX-LOOP-HELIX MOTIF.

```

FT CONFLICT 149 149 G -> A (IN REF. 1).
 SQ SEQUENCE 312 AA; 34116 MW; 069479287438F456 CRC64;

Query Match 11.4%; Score 213; DB 1; Length 312;
 Best Local Similarity 34.0%; Pred. No. 1e-06;
 Matches 68; Conservative 21; Mismatches 85; Indels 26; Gaps 6;

QY 33 PPQPAPA-----TLQAREHVPVPELSLLDSTD--RAWLAPTLLQGITARA 77
 DB 123 PVASPPAVEMGSSNVGCTKTPASAGP--KPRSVYTKNQKQSTTATSTPTTAAESSASV 180
 QY 78 AOYLLHSPELGASEAAAPRDEVDGRLVRRSSGGASSKSPQPKVREQLCKLKGVVV 137
 DB 181 NLYTEFFQNFDFNSALFDDSEDDLM-LFSGGEDFDGDSF-----DLADGENQ 232
 QY 138 DELGCSQRPASQKQVNG-VQKORLAANARERRMHGLNHFADQLRVNVPISFNNKKLS 196
 DB 233 DAAAGSGGKRRGKQITPVVKKRRLAANARERRRMQNLQAFDLRLQYLPCLGNDRQLS 292
 QY 197 KYETLQMAQIYINALSLLQ 216
 DB 293 KHETLQMAQIYISALGDLRL 312

RESULT 9
 ID NDF2 HUMAN STANDARD; PRT; 382 AA.
 AC Q15784;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 2 (NeuroD2).
 GN NEUROD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413331; PubMed=8816493;
 RA McCormick M.B., Tamimi R.M., Snider L., Asakura A., Bergstrom D.,
 RA Tapscott S.J.;
 RT "NeuroD2 and neuroD3: distinct expression patterns and
 RT transcriptional activation potentials within the neuroD gene
 RT family.";
 RL Mol. Cell. Biol. 16:5792-5800 (1996).
 RN [2]
 RP REVISIONS TO 336-338.
 RA Tapscott S.J., Tamimi R.T., McCormick B.M.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Appears to mediate neuronal differentiation.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U58681; AAC26058.1; -.
 DR TRANSFAC; T04903; -.
 DR Genew; HGNC:7763; NEUROD2.
 DR MIM; 601725; -.
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR GO; GO:0007399; P:neurogenesis; TAS.
 DR GO; GO:0006357; P:regulation of transcription from Pol II promoter; TAS.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.
 DR PROSITE; PS00888; HLH; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Differentiation.
 FT DOMAIN 37 46 POLY-PRO.
 FT DOMAIN 82 91 POLY-GLU.
 FT DOMAIN 107 113 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DNA_BIND 122 133 BASIC DOMAIN.
 FT DOMAIN 134 174 HELIX-LOOP-HELIX MOTIF.
 FT DOMAIN 174 285 POLY-GLY.
 SQ SEQUENCE 382 AA; 41361 MW; E33AA5C0BE371B47 CRC64;

Query Match 11.3%; Score 210; DB 1; Length 382;
 Best Local Similarity 26.6%; Pred. No. 1.9e-06;
 Matches 98; Conservative 41; Mismatches 119; Indels 110; Gaps 17;

QY 22 QPQHLLPQPPPPPPATLQAREHVPVPELSLLDSTDPRAWLAPTLLQGITARAQYL 81
 DB 28 EPRSDKGDAPPPPPAP-----GPGAGP--PARAAKPV 58
 QY 82 LHSPELGASEAAAPRDEVDGRLVRRSSGGASSKSPQPKVREQLCKLKGVVVDELG 141
 DB 59 ---PLRGEETATLAETKEGEL-----GGEEREE---EKEGLDAEG-----E 99
 QY 142 CSQRAPSSQVNGVQ---KORLAANARERRMHGLNHFADQLRVNVPISFNNKKLSK 197
 DB 100 RPKRGPKKRMKTKARLERSKLRRQKANARERRMHDLNALDNLKRVVPCYSKTKLSK 159
 QY 198 YETLQMAQIYINALSLL-----QTPSGEQPPPP--PASC-KSDHHHLRTAA 242
 DB 160 IETLRKNIWALSSEILRSKRPDLVSVVQTLCKGLSQFTTNLVAGCLQLSRNLFTQ 219
 QY 243 SYEG-----GAGNATA-----AGAQ-OASGG-----SQRPPTPGSCETRESAPA 280
 DB 220 GADGAGRFGHSGGPFAMHPYPCSLAGAACQAGGLGGGAHALRTGYCAAVETLVA 279
 QY 281 SAGYSVQLDALHFTSFEDSALTAMAAKNLSPSLPGSILQPVQENSKTSRSHRS DGE 340
 DB 280 AAGGGGASPD-VNSSEYEGPLPPLCLNGNFS-----LKQDSSPDHEKSY----- 323
 QY 341 FSPHSHYS 348
 DB 324 -----HYS 326

RESULT 10
 NGN1 HUMAN STANDARD; PRT; 237 AA.
 ID NGN1 HUMAN STANDARD; PRT; 237 AA.
 AC Q92886; Q96HE1;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
 DE (Neurogenic basic-helix-loop-helix protein).
 GN NEUROG1 OR NGN1 OR NGN OR NEUROD3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413331; PubMed=8816493;
 RA McCormick M.B., Tamimi R.M., Snider L., Asakura A., Bergstrom D.,
 RA Tapscott S.J.;
 RT "NeuroD2 and neuroD3: distinct expression patterns and
 RT transcriptional activation potentials within the neuroD gene
 RT family.";
 RL Mol. Cell. Biol. 16:5792-5800 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA R. Richards S., Worley K.C., Haje S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Appears to mediate neuronal differentiation.
CC -1- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expression restricted to the embryonic
CC nervous system.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U63842; AAB37575.1; -.
CC DR EMBL: BC008687; AAH08687.1; -.
CC DR EMBL: BC028226; AAH28226.1; -.
CC DR TRAFANGAC; T04907; -.
CC DR Genew; HGNC:7764; NEUROG1.
CC DR MIM; 601726; -.
CC DR GO; 0003700; F:transcription factor activity; TAS.
CC DR GO; 0007399; P:neurogenesis; TAS.
CC DR GO; 0006357; P:regulation of transcription from Pol II pro. .; TAS.
CC DR InterPro; IPR001092; HLH_basic.
CC DR Pfam; PF00010; HLH; 1.
CC DR SMART; SM00353; HLH; 1.
CC DR PROSITE; PS00888; HLH; 1.
CC DR DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Neurogenesis; Developmental protein; Differentiation.
CC FT DNA BIND 93 104 BASIC DOMAIN.
CC FT DOMAIN 105 145 HELIX-LOOP-HELIX MOTIF.
CC FT CONFLICT 51 51 G->S (IN REF.1).
CC SQ SEQUENCE 237 AA; 25717 MW; E370A0B3408D567 CRC64;
Query Match 11.3%; Score 209.5; DB 1; Length 237;
Best Local Similarity 33.8%; Pred. No. 1.2e-06;
Matches 70; Conservative 26; Mismatches 64; Indels 47; Gaps 9;
QY 107 RRSFGAS-SKSPGVKVEQICKLKGVVVDLGCGRAPSPKSGKGVGVQKORLAAN 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 49 REGAPNISRASVEGAQDDQERRRRG-----RTRVRSEALLSLRSRYKAN 98
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 ARRRRHGHLNHPDOLRNVPSPFNNDKLSKVETLQMAQIYINALSELQ-----TPSG 220
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 99 DRENRMHNLNMAIDALRSVLPFPDDTKLTKETLRFANVYIWAETLRLADQLPGG 158
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 GEQ-----PPPPASCKSDHHLRTAASVEGGAGNATAAGAAQAGSGSQ-----RP 266
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 159 GARERLLPPQCVCLPGPPSPASD-----AESWGSGAAAASPLSDPSPAAASEDFIYRP 212
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 TPQSCCTRPSAPASAGYSVQDALH 293
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 277 SAPASAGGYVOLDALHFSTEDSALTAMMAQKNLSPLSPQLPVQ----- 324
Db 248 -----YSATLE-----PFESPUDC-----TSFSPDGLSPPLSINGNFSFKHEPS 289
QY 325 ---EEN-----SKTSPSRHS-----DG--EFSPHSY 347
Db 290 AEFKNYAFTHYPAATLAGPQSHGIFSGATAPRCEIPIDNIMSPDSSH 341

RESULT 12
ID NDF6 MOUSE STANDARD; PRT; 337 AA.
AC P48986;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenic differentiation factor 6 (NeuroD6) (Atonal protein homolog
2) (Helix-loop-helix protein MATH-2) (MATH2) (NEX-1 protein).
GN NEUROD6 OR ATOH2 OR ATH2 OR NEX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RX MEDLINE=95262673; PubMed=7744035;
RA Shimizu C., Akazawa C., Nakanishi S., Kageyama R.;
RT "MATH-2, a mammalian helix-loop-helix factor structurally related to
the product of Drosophila proneural gene atonal, is specifically
expressed in the nervous system.";
RL Eur. J. Biochem. 229:239-248(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=95200803; PubMed=7545978;
RA Bartholomae A., Nave K.-A.;
RT "NEX-1: a novel brain-specific helix-loop-helix protein with
autoregulation and sustained expression in mature cortical neurons.";
RL Mech. Dev. 48:217-228(1994).
CC -!- FUNCTION: Activates E box-dependent transcription in collaboration
with E47. May be a trans-acting factor involved in the development
and maintenance of the mammalian nervous system. Transactivates
the promoter of its own gene.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Specific to the nervous system of both embryos
and adults. Highest levels in the cortical plate of the cerebrum.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D44480; BAA07923.1; -.
DR EMBL; U29086; AAC14576.1; -.
DR PIR; I48682; I48682.
DR PIR; I57038; I57038.
DR MGD; MGI:106593; Neurod6.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DOMAIN 54 63 POLY-GLU.
FT DOMAIN 80 86 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DNA BIND 95 106 BASIC DOMAIN.
FT DOMAIN 107 147 HELIX-LOOP-HELIX MOTIF.
SQ SEQUENCE 337 AA; 38644 MW; 35C18ACD8E1BFBA CRC64;
Query Match 11.0%; Score 205.5; DB 1; Length 337;
Best Local Similarity 29.0%; Pred. No. 3.2e-06;
Matches 71; Conservative 31; Mismatches 94; Indels 49; Gaps 7;
QY 102 RGEIVRRSSGGASSKSPGVKVRQQLCKLKGVVVDLGCGRQRAPSPSKQVNGVQ--- 157
Db 44 RKSIKRAFGTEETEKEEEDREEE-----DENGLRRRLKTKTKLRLERV 92
QY 158 KORLAANARERRRMHGLNHFQDLNRNVPFNNNDKLSKYETLQMAQIYINALSELLQT 217
Db 93 KFRQEANARERNMHLNDALDNLRKVKVPCYSKTKLKIETLRKAKNVIWALSILRI 152
QY 218 PSGGEOP-----PPPPASCKSDHHLRTAAASYEGGAGNATAAGAAQASGGSQ 264
Db 153 ---GKRPDLLTFTVQNLCKGLSQFTTNLVAGIQL-NARSLFMGGGGEAAHHTSPYSTFY 208
QY 265 RP-----TPPGSCRTRESAPASAGGYSVQLDALHFTFEDSALTAMMAQKNLSPLPG 317
Db 209 PPHSGPELATPPGHGTLDNSKMKPNYCSAYESFYESTSPECA-----SQPFEG 258
QY 318 SILQP 322
Db 259 PLSPPP 263

RESULT 13
ID NGN1 RAT STANDARD; PRT; 244 AA.
AC P70595;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
DE (Neurogenic basic-helix-loop-helix protein).
GN NEUROG1 OR NGN1 OR NGN OR NEUROD3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Embryo;
RX MEDLINE=97011137; PubMed=8858147;
RA Ma Q., Kintner C., Anderson D.J.;
RT "Identification of neurogenin, a vertebrate neuronal determination
gene.";
RL Cell 87:43-52(1996).
CC -!- FUNCTION: Appears to mediate neuronal differentiation.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
bHLH protein.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expression restricted to the embryonic
nervous system.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67777; AAC52857.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
KW Neurogenesis; Developmental protein; Differentiation.
FT DNA BIND 94 105 BASIC DOMAIN.

```

FT DOMAIN 106 146 HELIX-LOOP-HELIX MOTIF.
 SQ SEQUENCE 244 AA; 26182 MW; 47D0E91403143D61 CRC64;
 Query Match 11.0%; Score 205; DB 1; Length 244;
 Best Local Similarity 34.2%; Pred. No. 2.4e-06;
 Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;
 QY 95 PRDEVDGRGELVRESSGAS-SSKSPGVKVRQELCKLKGVVVDLGCSPQAPSSQV 153
 DB 38 ELASTGLSVPAARSAPTLGASNVPGQDEEORRRRG-----RARVRSALL 87
 QY 154 NGVOKORLAANARERRRMHGLNHFADQLRNVIPSFNNDKKLSKYETLQMAQIYINALSE 213
 DB 88 HSLRRSRVRKANDEERNMHNALDARSVLSPFPDDTKLTETLRFAYNYIWLAE 147
 QY 214 LLQ-----TSGGSEQ-PP-----PPPASCSDHHLHRTAASYGGGAGNATAAGAAQ 258
 DB 148 TRLADQGLPGGARERLLPPQCVCLPGPPSPASD-----TESWGGAAASPQATV 199
 QY 259 ASGGSORPTPPGS 271
 DB 200 ASPLSDPSSPSAS 212

RESULT 14
 NDF2 RAT
 ID NDF2 RAT STANDARD; PRT; 382 AA.
 AC Q63689;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic differentiation factor 2 (NeuroD2) (Brain bHLH protein
 DE KW8).
 GN NEUROD2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=96193685; PubMed=8605021;
 RA Kume H., Maruyama K., Tomita T., Iwatsubo T., Saido T.C., Obata K.;
 RT "Molecular cloning of a novel basic helix-loop-helix protein from the
 RT rat brain."
 RL Biochem. Biophys. Res. Commun. 219:526-530(1996).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Hulo C.;
 RL Unpublished observations (AUG-1998).
 CC -!- FUNCTION: Appears to mediate neuronal differentiation.
 CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
 CC bHLH protein.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Detected only in neural tissue.
 CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
 CC -!- CAUTION: This is a conceptual translation; two frameshifts were
 CC introduced in positions 225 and 250 to extend the similarity with
 CC the human and mouse orthologs.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D82868; BAAL1615.1; ALT FRAME.
 CC InterPro; IPR001092; HLH_basic.
 CC Pfam; PF00010; HLH; 1.
 CC SMART; SM00353; HLH; 1.
 CC PROSITE; PS00888; HLH; 1.

KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Differentiation.
 FT DOMAIN 82 91 POLY-GLU.
 FT DNA_BIND 122 133 BASIC DOMAIN.
 FT DOMAIN 134 174 HELIX-LOOP-HELIX MOTIF.
 FT DOMAIN 182 285 POLY-GLY.
 SQ SEQUENCE 382 AA; 41466 MW; 05214DE9BE0A8069 CRC64;
 Query Match 11.0%; Score 204; DB 1; Length 382;
 Best Local Similarity 23.5%; Pred. No. 4.6e-06;
 Matches 93; Conservative 32; Mismatches 89; Indels 182; Gaps 16;
 QY 18 DHRQPOPHLPPPPPP-----QPPATLQAREHPVYPPELSILDSTDPRAWLAP 67
 DB 27 DEPRSDKGDAPPQPPPPPPAPGSGAPGAPATKPVSLRGEVP-----EP 68
 QY 68 TLOGICTTAAAYLLHSPELGASFAAAPRDEVDGRGELVRRSSGSGASSKSPGVKVRQ 127
 DB 69 TLAEV-----KEEGELGGEEEEEE-EGLE-----AEGERPKKRGPCKRMT 112
 QY 128 LCKLKGVVVDLGCSPQAPSSQVNGVOKORLAANARERRRMHGLNHFADQLRNVIP 187
 DB 113 KARLER-----SKLRQKANARERNRMDLNALDNRKVP 149
 QY 188 SFNNKKLSKYETLQMAQIYINALSELLQ----- 217
 DB 150 CYSKTQKLSKETLRLAKNYIWLSEILRSKRPDLVSYVYVTLCKGLSQPTTNLVAGCLQ 209
 QY 218 -----PSGGE-----QPPPPPPAS-----CKS-----DHHLRT-- 240
 DB 210 LNSRNFLTEQGADGAGRFHSGGGPFAMHPYPYPCSLAGDQCOAAGGLGGGAHALRTHG 269
 QY 241 -AASYE-----GGAGNATAAGAQSGGSGQRPTPGSCRTRESAPASAGYSVQLD--- 290
 DB 270 YCAAYETLYAAGGGGASPDYNSVEYEG-----PLSPPLCLN-----GNFSLKQDSSP 317
 QY 291 -----ALHFTTFEDSALTAMMAQKNLSPLSPGS 318
 DB 318 DHEKSYHYMSHS-----AUPGS 335
 RESULT 15
 NG1 MOUSE
 ID NG1 MOUSE STANDARD; PRT; 244 AA.
 AC P70660;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenin 1 (Neurogenic differentiation factor 3) (NeuroD3)
 DE (Neurogenic basic-helix-loop-helix protein) (Helix-loop-helix protein
 DE MATH-4C) (MATH4C).
 GN NEUROG1 OR NGN1 OR NGN OR NEUROD3 OR ATH4C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=96413331; PubMed=8816493;
 RA McCormick M.B., Tammi R.M., Snider L., Asakura A., Bergstrom D.,
 RA Tapscott S.J.;
 RT "NeuroD2 and NeuroD3: distinct expression patterns and
 RT transcriptional activation potentials within the neuroD gene
 RT family."
 RL Mol. Cell. Biol. 16:5792-5800(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RX MEDLINE=97011137; PubMed=8858147;
 RA Ma Q., Kintner C., Anderson D.J.;
 RT "Identification of neurogenin, a vertebrate neuronal determination
 RT gene.";

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 21, 2004, 21:04:52 ; Search time 136.88 Seconds
(without alignments)
815.995 Million cell updates/sec

Title: US-09-980-381a-58
Perfect score: 1862
Sequence: 1 MSRLHAEWAEEVKEGLGDH.....HRSDGEFSPHYSDSEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582	31.3	266	13	Q42245
2	532	28.6	177	13	Q800U5
3	337	18.1	161	13	Q98939
4	309	16.6	64	13	Q8JHA1
5	232.5	12.5	255	5	Q9NE27
6	228.5	12.3	178	13	Q9YH81
7	228.5	12.3	179	13	Q9PWP1
8	224.5	12.1	213	13	Q9YH82
9	223	12.0	211	13	Q9PWP2
10	222.5	11.9	316	13	Q9W6C7
11	222	11.9	272	4	Q8N416
12	220.5	11.8	213	13	Q9PTR9
13	219.5	11.8	347	13	Q9DE43
14	218.5	11.7	347	13	Q8JH35
15	217.5	11.7	195	5	Q9VGJ5
16	217	11.7	151	13	O57598

```

17 212.5 11.4 189 5 Q9XZC7
18 211.5 11.4 325 13 Q9W6C6
19 210.5 11.3 381 4 Q9UCQ6
20 210.5 11.3 382 4 Q8TB17
21 210 11.3 350 13 Q42202
22 209 11.2 134 13 Q8JHA0
23 207 11.1 138 13 O13126
24 206 11.1 197 11 P70562
25 204.5 11.0 220 13 Q8JIS0
26 204 11.0 214 11 O08718
27 202.5 10.9 189 5 Q9V7M2
28 202.5 10.9 189 5 Q9V7M2
29 202 10.8 152 4 Q8N100
30 202 10.8 197 11 Q9QVC3
31 202 10.8 357 11 Q8CE17
32 201.5 10.8 337 4 Q81YR9
33 201 10.8 149 11 Q9Z2E5
34 200 10.7 216 13 Q9W6B8
35 199.5 10.7 337 4 Q7Z4W9
36 197 10.6 134 13 Q9DGA9
37 197 10.6 134 13 Q8AW52
38 197 10.6 356 4 Q9UEC8
39 196.5 10.6 214 11 Q9DB10
40 196.5 10.6 214 13 P70050
41 194 10.4 309 13 Q919C9
42 193 10.4 46 13 Q88904
43 192 10.3 336 13 Q9DDQ8
44 192 10.3 390 5 Q9VID2
45 189 10.2 355 11 Q9JL05

```

ALIGNMENTS

RESULT 1

```

O42245 ID O42245 PRELIMINARY; PRT; 266 AA.
AC O42245;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Atonal homologue-1.
GN ATOH1 OR ZATH-1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim C.-H., Bae Y.-K., Yamanaka Y., Yamashita S., Shimizu T., Fujii R.,
RA Hibl M., Hirano T.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF024536; AAB82272.1; -.
DR ZFIN; ZDB-GENE-950415-17; atoh1.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
SQ SEQUENCE 266 AA; 28786 MW; 56A92494B52PEFB8 CRC64;

```

Query Match 31.3%; Score 582; DB 13; Length 266;

Best Local Similarity 46.5%; Pred. No. 1e-38;

Matches 155; Conservative 21; Mismatches 69; Indels 88; Gaps 14;

QY 12 EVKELGDHRRQPQPHLPQPPPPPPPPATLQAREHPVVPPELSLSDTDPRAWLAPTQ 71

DB 10 EVVELDVQH-----SSLGEGESKYPPALALNASSDPRAWLAPVQAG 51

QY 72 ICTAPAAQYLLHSPELGASEAAAPRDEYDGRGELVRRSSGGASSKSPGPVKVRQLCKL 131

```
Db 52 TCAAH-AEYLLHSP--GSS-----AECVSSASNFRKSSKS--PVKVR-LCRL 93
Qy 132 KGGVVVDELGCSORAPSSKQVNGOKORLAANARERRMHGLNHAFDOLRNVIPIFNN 191
Db 94 KGAVGADE---GQRAPSSKSTNVQKORMAANARERRMHGLNHAFDELRSVIPAFDN 150
Qy 192 DKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHH-LRTAAAYEGGAGN 250
Db 151 DKLSKYETLQMAQIYINALSDLLQPGAKADPP-----NCDLLHANVLETRPRGS--- 203
Qy 251 ATAAGAAQAGSGQRPTPGSCRTFSAPASAGYSVQLDALHFSTFEDSALTAMMAQKN 310
Db 204 -----PGVCR-----GTGVGYPYQ-----YEDGTFSFMEQDL 232
Qy 311 LSP---SLPGSILOPVOEENSKTSPRSHRSDGE 340
Db 233 QSPSGTSKSGS-----EASKDPSRNSRSDGE 258

RESULT 2
Q800U5 PRELIMINARY; PRT; 177 AA.
AC Q800U5;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE CATH1 (Fragment).
GN CATH1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert P.J., Timmer J.R., Helms A.W., Nakada Y., Hunsaker T.L.,
RA Johnson J.E.;
RT "Control of Math1 Expression by Interactions of Zic1 and Ebf2 with the
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP467292; AAO59913.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 177 AA; 19170 MW; FD44269C5994F5E9 CRC64;

Query Match 28.6%; Score 532; DB 13; Length 177;
Best Local Similarity 59.6%; Pred. No. 6.1e-35;
Matches 121; Conservative 12; Mismatches 44; Indels 26; Gaps 5;

Qy 152 QVNGVOKORLAANARERRMHGLNHAFDOLRNVIPIFNNDKLSKYETLQMAQIYINAL 211
Db 1 QVSGVOKORLAANARERRMHGLNHAFDOLRNVIPIFNNDKLSKYETLQMAQIYISAL 60
Qy 212 SELLQTPSGGEQPPPPASCKSDHHHRLTAASVEGGAGNATAAGAAQAGSGSORPTPGS 271
Db 61 AELLHGPPA--PPEPPAKAE-----LR-GAPF-----PPPPPPPPPPAP 102
Qy 272 CTRFSAPASAGYSVQLDALHFSTFEDSALTAMMAQKNLSPSLPGSILOPVOEENSKTS 331
Db 103 ARTFRPAAAGGFAAQLEPLRPPSPF-----AQKAPSPALLGPPAPQPPERSKAS 154
Qy 332 PRSHRSDGFFSHSHVSDSEAS 354
Db 155 PRSHRSDGFFSHSHVSDSEAS 177

RESULT 3
Q98939 PRELIMINARY; PRT; 161 AA.
ID Q98939
AC Q98939;
```

```
DT 01-FEB-1997 (TEMBLrel. 02, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE CATH1 (Fragment).
GN CATH1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben-Arie N., McCall A.E., Berkman S., Eichele G., Bellen H.J.,
RA Zoghbi H.Y.;
RT "Evolutionary conservation of sequence and expression of the BHLH
RT protein Atonal suggests a conserved role in neurogenesis.";
RL Hum. Mol. Genet. 5:1207-1216(1996).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; U61149; AAB41304.1; -.
DR TRANSFAC; T04545; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS50888; HLH 2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 161 AA; 16652 MW; BD0B73B5AA3F9F67 CRC64;

Query Match 18.1%; Score 337; DB 13; Length 161;
Best Local Similarity 61.2%; Pred. No. 2.4e-19;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

Qy 107 RRSGGAS---SSKSPGVKVRQLCKLKGVVVDLGCGRQRPASCKQVGVQKORRLA 163
Db 42 RPVAGGAARVPAGAPGE---RGAAGARGGG---GGAGPRA---QVSGVQKORRLA 89
Qy 164 ANARERRRMHGLNHAFDOLRNVIPIFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQ 223
Db 90 ANARERRRMHGLNHAFDOLRNVIPIFNNDKLSKYETLQMAQIYISALAEHLHGPPA--- 146
Qy 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 4
Q8JHA1 PRELIMINARY; PRT; 64 AA.
AC Q8JHA1;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE BHLH transcription factor ath1 (Fragment).
OS Serinus canaria (Canary).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC Fringillidae; Carduelinae; Serinus.
NCBI_TaxID=9135;
RN [1]
RP SEQUENCE FROM N.A.
RA Huverstuhl J., Brors D., Bodmer D., Mullen L., Gleich O., Strutz J.,
RA Ryan A.F.;
RT "Expression of BHLH and class IV POU-domain transcription factors in
RT the chicken and canary inner ear.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL; AF526421; AAM89247.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
```

```

RESULT 6
Q9YH81
ID Q9YH81 PRELIMINARY; PRT; 178 AA.
AC Q9YH81;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Neurogenin 1.
GN NGN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=2106448; PubMed=11124117;
RA Matter-Sadzinski L., Matter J.M., Ong M.T., Hernandez J., Ballivet M.;
RT "Specification of neurotransmitter receptor identity in developing
RT retina: the chick ATH5 promoter integrates the positive and negative
RT effects of several bHLH proteins.";
RL Development 128:217-231 (2001).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR EMBL: AJ012660; CAA10106.1; -.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS50888; HLH 2; 1.
SQ SEQUENCE 178 AA; 19058 MW; 7EBCC1CD1ADA5633 CRC64;

Query Match 12.3%; Score 228.5; DB 13; Length 178;
Best Local Similarity 34.7%; Pred. No. 1.4e-10;
Matches 77; Conservative 26; Mismatches 54; Indels 65; Gaps 11

Qy 109 SSGGASSKSPGPVKVREQLCKLKGVVVDLGGSRQAPSSKQVGVQKORRLAANARE 168
Db 7 SSGGV--SEPPGAPRR-----RRRG-----RARPTEALLHTLKRSPRVKANDRE 51

Qy 169 RRRMHGHNHAFDQLNRNVPISFNNDKLSKYETLOWAQIYINALSLLQTPSGGEQ--PPP 226
Db 52 RNRMHNLNALDELRSVLPTFPDDTKLTKIETLRFAYNIWALSETLRL--AEQCLPPP 108

Qy 227 P-----PASCCKDHHHLRTAA SYEGGAGNATAAGAAQOAGGSGORPPPGSCRTFSAP 279
Db 109 PAFRGPAPPSPGSD-----AGSWLSGSGS---PAAPSLC-----139

Qy 280 ASAGYSVQLDALHFSFEDSALT---AMMAOKNLSPLPGS 318
Db 140 ASAGSPSP-----ATSEDCGYPSDALRFRGLPPAAGA 175

RESULT 7
Q9PWF1
ID Q9PWF1 PRELIMINARY; PRT; 179 AA.
AC Q9PWF1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Neurogenin 1.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=99180576; PubMed=10079233;
RA Perez S.E., Rebelo S., Anderson D.J.;
RT "Early specification of sensory neuron fate revealed by expression and
RT function of neurogenins in the chick embryo.";

```

RL	Development 126:1715-1728(1999).
CC	-!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.
DR	EMBL; AF123883; AAD22059.1; -.
DR	InterPro; IPR001092; HLH_basic.
DR	Pfam; PF00010; HLH; 1.
DR	SMART; SM00353; HLH; 1.
DR	PROSITE; PS00888; HLH_2; 1.
SQ	SEQUENCE 179 AA; 19189 MW; 1EF1F509869D91 CRC64;
Query Match 12.3%; Score 228.5; DB 13; Length 179;	
Best Local Similarity 34.7%; Pred.No.1.4e-10;	
Matches 77; Conservative 26; Mismatches 54; Indels 65; Gaps 11;	
QY	109 SSGASSSKSPGVKVEQCKLKGVVVDLGCGRAPSSKWVGQVKORLLAANARE 168 : : : : : : : : : : : : : : : : : : : : 8 SSGGV-SEPPGAPRRR---RRRG-----RARARTEALLHTLKSRRYKVANDRE 52
Db	
QY	169 RRRWHGLNHFADQLRNVPISFNNDKKLSKYETLQMAQIYNALSELLQTSGGEQ---PPP 226 : : : : : : : : : : : : : : : : : : : : 53 RNRMHLLNALDLRSVLPFPDDTKLTKTETLFAYNIWALSETLRLL---AQCLPPP 109
Db	
QY	227 P-----PASCSDHHHLRTAAIEGGAGNATAAGQAQSGSQRPDPSCSCHTRFSAP 279 : : : : : : : : : : : : : : : : : : : : 110 PAFRGPAPPSPGSD-----AGSWLSSGS---PAAPSLC----- 140
Db	
QY	280 ASAGGYSVQDALHFFSTFEDSALT---AMMAQNKLSPSLPGS 318 : : : : : : : : : : : : : : : : : : : : 141 ASAGPSPP-----ATSDCGYVPSDALRAFRGLPFAAFGA 176
Db	
RESULT 8	
ID QYH82	PRELIMINARY; PRT; 213 AA.
AC QYH82;	
DT 01-MAY-1999	(TtEMBLrel. 10, Created)
DT 01-MAY-1999	(TtEMBLrel. 10, Last sequence update)
DT 01-MAR-2003	(TtEMBLrel. 23, Last annotation update)
DE Neurogenin 2,	
GN NGN2.	
OS Gallus gallus (Chicken).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
OC Gallus.	
OX NCBI_TaxID=9031;	
RX [1]	
RP SEQUENCE FROM N.A.	
RA Ballivet M., Alliod C., Matter J.M.;	
RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.	
RN [2]	
RP SEQUENCE FROM N.A.	
RA Simmons A.D., Horton S., Abney A.L., Johnson J.E.;	
RT "Neurogenin 2 Expression in Ventral and Dorsal Spinal Neural Tube Progenitor Cells is Regulated by Distinct Enhancers.";	
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.	
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF TRANSCRIPTION FACTORS.	
CC EMBL; AJ012659; CAAL0105.1; -.	
DR EMBL; AF303000; AAG40768.1; -.	
DR InterPro; IPR001092; HLH_basic.	
DR Pfam; PF00010; HLH; 1.	
DR SMART; SM00353; HLH; 1.	
DR PROSITE; PSS0888; HLH_2; 1.	
SQ SEQUENCE 213 AA; 22507 MW; 24036E2C9A25251A CRC64;	
Query Match 12.1%; Score 224.5; DB 13; Length 213;	
Best Local Similarity 29.7%; Pred.No.3.7e-10;	
Matches 78; Conservative 24; Mismatches 88; Indels 73; Gaps 9;	
QY	32 PPPQPPTLAQREHPVYPPELSLDSTDFAWLAPTQCICTARAQYLHSPELGASE 91 : : : : : : : : : : : : : : : : : : : : 2 PVKAESPA-----PAADELLHRIASP---AP-----SASTDSSGA 24
Db	

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 20:34:21 : Search time 167,088 Seconds
(without alignments)
598.618 Million cell updates/sec

Title: US-09-980-381A-58
Perfect score: 1862
Sequence: 1 MSRLHAEWAWEVKELGDH.....HRSDGEFSPHSHYSDSEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1862	100.0	354	4 AAB60375	Aab60375 Human ato
2	1862	100.0	354	4 AAB60349	Aab60349 Human ato
3	1862	100.0	354	6 AAO30945	Aao30945 Human ato
4	1862	100.0	354	6 AAO30920	Aao30920 Human ato
5	1627.5	87.4	351	4 AAB60353	Aab60353 Mouse ato
6	1627.5	87.4	351	4 AAB60369	Aab60369 Mouse ato
7	1627.5	87.4	351	6 AAO30939	Aao30939 Mouse ato
8	1627.5	87.4	351	6 AAO30924	Aao30924 Mouse ato
9	582	31.3	266	4 AAB60366	Aab60366 Zebrafish
10	582	31.3	266	6 AAO30936	Aao30936 Zebra fis
11	537	28.8	259	6 AAO30949	Aao30949 Frog aton
12	337	18.1	161	4 AAB60376	Aab60376 Chicken a
13	337	18.1	161	4 AAO30946	Aao30946 Chicken a
14	228.5	12.3	178	4 AAB60356	Aab60356 Chicken a
15	228.5	12.3	178	6 AAO30926	Aao30926 Chicken a
16	228.5	12.3	198	4 ABB71919	Abb71919 Drosophil
17	224.5	12.1	213	4 AAB60357	Aab60357 Chicken a
18	224.5	12.1	213	6 AAO30927	Aao30927 Chicken a
19	222	11.9	250	5 AAE29281	Aae29281 Human neu
20	222	11.9	272	7 ADC31349	Adc31349 Human nov
21	218	11.7	263	4 AAB60351	Aab60351 Mouse ato
22	218	11.7	263	4 AAB60365	Aab60365 Mouse ato
23	218	11.7	263	6 AAO30935	Aao30935 Mouse ato
24	218	11.7	263	6 AAO30922	Aao30922 Mouse ato
25	217.5	11.7	195	4 ABB62789	Abb62789 Drosophil

ALIGNMENTS

RESULT 1

AAB60375
ID AAB60375 standard; protein; 354 AA.

XX AC AAB60375;

DT 24-APR-2001 (first entry)

DE Human atonal homologue 1 (Hathl) protein, SEQ ID NO:58.

XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.

XX OS Homo sapiens.

XX PN WO200073764-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015410.

XX 01-JUN-1999; 99US-0137060P.

XX 19-JAN-2000; 2000US-0176993P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;

XX WPI; 2001-032190/04.

XX N-PSDB; AAF27283.

XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any of its homologues or orthologs, for the treatment of e.g. deafness, osteoarthritis and abnormal cell proliferation.

XX Claim 34; Page; 142pp; English.

XX The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated

Aab60364 Chicken a
Aao30934 Chicken a
Aaw54946 Mouse neu
Aay70569 Murine ne
Aab60362 Mouse neu
Aao30932 Mouse ato
Aar77505 Frog neur
Aaw22437 Frog neur
Aaw71010 Xenopus n
Abg71796 Frog bHLH
Aae29280 Human neu
Abb63330 Drosophil
Aab60378 Drosophil
Aao30948 Drosophil
Aab60355 Drosophil
Aao30925 Drosophil
Aaw22439 Human neu
Aaw71015 Human neu
Abg72001 Human bHL
Aab14347 Human neu

26 217 11.7 151 4 AAB60364
27 217 11.7 151 6 AAO30934
28 217 11.7 263 2 AAW54946
29 217 11.7 263 3 AAY70569
30 217 11.7 263 4 AAB60362
31 217 11.7 263 6 AAO30932
32 214 11.5 352 2 AAR77505
33 214 11.5 352 2 AAW22437
34 214 11.5 352 2 AAW71010
35 214 11.5 352 6 ABG71796
36 213.5 11.5 237 5 AAE29280
37 213 11.4 312 4 ABB63330
38 213 11.4 312 4 AAB60378
39 213 11.4 312 6 AAO30948
40 212.5 11.4 189 4 AAB60355
41 212.5 11.4 189 6 AAO30925
42 210.5 11.3 381 2 AAW22439
43 210.5 11.3 381 2 AAW71015
44 210.5 11.3 381 6 ABG72001
45 210 11.3 382 3 AAB14347

CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank
 XX
 SQ Sequence 354 AA;

Query Match 100.0%; Score 1862; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAELGKELGHHQPPHLPQPPPPPPATLQAREHVPVPELSDSTD 60
 Db 1 MSRLHAEWAELGKELGHHQPPHLPQPPPPPPATLQAREHVPVPELSDSTD 60
 QY 61 PRAWLPTLOGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 Db 61 PRAWLPTLOGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 QY 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHA 180
 Db 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHA 180
 QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGAAQAGSGGSRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 Db 241 AASYEGAGNATAAGAAQAGSGGSRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 QY 301 ALTAMMAQKNLSPLSGSILQPVQENSKTSPRSHRSDGFEFPHSHYSVSDSDEAS 354
 Db 301 ALTAMMAQKNLSPLSGSILQPVQENSKTSPRSHRSDGFEFPHSHYSVSDSDEAS 354

RESULT 2
 ID AAB60349
 XX AAB60349 standard; protein; 354 AA.
 AC AAB60349;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human atonal homologue 1 (ATOH1, Hath1) protein, SEQ ID NO:2.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.
 XX
 OS Homo sapiens.
 XX
 FN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HV, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 PI

XX WPI: 2001-032190/04.
 DR N-PSDB; AAF27252.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologs or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank
 XX
 SQ Sequence 354 AA;

Query Match 100.0%; Score 1862; DB 4; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSRLHAEWAELGKELGHHQPPHLPQPPPPPPATLQAREHVPVPELSDSTD 60
 Db 1 MSRLHAEWAELGKELGHHQPPHLPQPPPPPPATLQAREHVPVPELSDSTD 60
 QY 61 PRAWLPTLOGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 Db 61 PRAWLPTLOGICTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 QY 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHA 180
 Db 121 PVKVRQOLCKLGGVVVDLGCSPORAPSSKQVNGVKORRLAANARERRMGLNHA 180
 QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELLQTPSGGEQPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGAAQAGSGGSRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 Db 241 AASYEGAGNATAAGAAQAGSGGSRPTPGSCRTFRSAPASAGYSVQLDALHFSTPDS 300
 QY 301 ALTAMMAQKNLSPLSGSILQPVQENSKTSPRSHRSDGFEFPHSHYSVSDSDEAS 354
 Db 301 ALTAMMAQKNLSPLSGSILQPVQENSKTSPRSHRSDGFEFPHSHYSVSDSDEAS 354

RESULT 3
 ID AAO30945
 XX AAO30945 standard; protein; 354 AA.
 AC AAO30945;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hath1) protein #2.
 XX
 KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW

KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hath1; human.
 XX
 OS Homo sapiens.
 XX WO2003047532-A2.
 PN
 XX 12-JUN-2003.
 PD
 XX
 PF 03-DEC-2002; 2002WO-US041458.
 XX
 PR 05-DEC-2001; 2001US-00004717.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HY, Qi Y;
 XX WPI; 2003-505253/47.
 DR N-PSDB; AAL62230.
 XX
 XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 XX Claim 33; Page 151-152; 157pp; English.
 PS
 XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hath1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 XX Sequence 354 AA;
 PS Query Match 100.0%; Score 1862; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MSRLHAEEWAEVKEVKGELGDHHRQPQPHLLPQPPPPPPPPATLQAREHPVYPPELSDST 60
 DB 1 MSRLHAEEWAEVKEVKGELGDHHRQPQPHLLPQPPPPPPPPATLQAREHPVYPPELSDST 60
 QY 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 DB 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 QY 121 PVKVRQLCKLKGVVVDDELGCSQRAPSSKQVNGVKORRLAANARERRMHLNHF 180
 DB 121 PVKVRQLCKLKGVVVDDELGCSQRAPSSKQVNGVKORRLAANARERRMHLNHF 180
 QY 181 QLRNVTPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGEGOPPPPPASCKSDHHHLRT 240
 DB 181 QLRNVTPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGEGOPPPPPASCKSDHHHLRT 240
 QY 241 AASYEGAGNATAAGQAQSGSGSQRTTPGSCRTFSAPASAGYSVQLDALHFSFEDS 300
 DB 241 AASYEGAGNATAAGQAQSGSGSQRTTPGSCRTFSAPASAGYSVQLDALHFSFEDS 300
 QY 301 ALTAMQAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354
 DB 301 ALTAMQAQKNLSPSLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354

RESULT 4
 AA030920
 ID AA030920 standard; protein; 354 AA.
 XX
 AC AA030920;
 XX
 DT 22-SEP-2003 (first entry)
 XX
 DE Human atonal homologue 1 (Hath1) protein #1.
 XX
 XX Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Hath1; human.
 XX
 OS Homo sapiens.
 XX WO2003047532-A2.
 PN
 XX 12-JUN-2003.
 PD
 XX
 PF 03-DEC-2002; 2002WO-US041458.
 XX
 PR 05-DEC-2001; 2001US-00004717.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HY, Qi Y;
 XX WPI; 2003-505253/47.
 DR N-PSDB; AAL62199.
 XX
 XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 XX Disclosure; Page 115-117; 157pp; English.
 PS
 XX The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result or loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is human atonal homologue 1 (Hath1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 XX Sequence 354 AA;
 PS Query Match 100.0%; Score 1862; DB 6; Length 354;
 Best Local Similarity 100.0%; Pred. No. 4.2e-151;
 Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MSRLHAEEWAEVKEVKGELGDHHRQPQPHLLPQPPPPPPPPATLQAREHPVYPPELSDST 60
 DB 1 MSRLHAEEWAEVKEVKGELGDHHRQPQPHLLPQPPPPPPPPATLQAREHPVYPPELSDST 60
 QY 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 DB 61 PRAWLAPTQIGICTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGASSKSPG 120
 QY 121 PVKVRQLCKLKGVVVDDELGCSQRAPSSKQVNGVKORRLAANARERRMHLNHF 180
 DB 121 PVKVRQLCKLKGVVVDDELGCSQRAPSSKQVNGVKORRLAANARERRMHLNHF 180

QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 Db 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 QY 241 AASVEGGAGNATAAGAAQAGSGSQRPPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 Db 241 AASVEGGAGNATAAGAAQAGSGSQRPPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 QY 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSFPHSHYSDSDEAS 354
 Db 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSFPHSHYSDSDEAS 354

RESULT 5
 AAB60353
 ID AAB60353 standard; protein; 351 AA.
 XX
 AC AAB60353;
 DT 24-APR-2001 (first entry)
 XX
 DE Mouse atonal homologue 1 (ATOH1, Math1) protein, SEQ ID NO:11.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.
 XX
 OS Mus musculus.
 XX
 PN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 XX
 DR WPI; 2001-032190/04.
 DR N-PSDB; AAF27257.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or amino acid sequence, or any of its homologues or orthologues as therapeutic agents for the treatment of deafness, partial hearing loss, vestibular effects due to damage or loss of inner hair cells, osteoarthritis and abnormal cell proliferation. The invention also encompasses methods of screening for compounds which affect the expression of an atonal-associated nucleic acid sequence in an animal, and a transgenic animal in which an allele of a native atonal-associated gene is replaced by a heterologous nucleic acid sequence, thus inactivating the atonal-associated allele. The nucleic acids or proteins may be used in a method of treating an animal for hearing impairment. CC joint disease, balance disorders, abnormal cell proliferation, or other disease related to loss of a functional atonal-associated nucleic acid or protein. They may particularly be used to treat an animal with a deficiency in cerebellar granule neurons or their precursors, and may also be used in promoting mechanoreceptive cell growth and generating hair cells. The present sequence represents an atonal-associated amino acid sequence referred to in the invention. Note: The present sequence is not shown in the specification, but was obtained from GenBank

XX SQ Sequence 351 AA;
 Query Match 87.4%; Score 1627.5; DB 4; Length 351;
 Best Local Similarity 88.4%; Pred. No. 5.4e-131;
 Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
 QY 1 MSRLHAEAEVAEVLGDDHRRQPOPHLPQPPPPPPATLQAREHPVYPPELSLDSTD 60
 Db 1 MSRLHAEAEVAEVLGDDHRRQPOPHV--PPLTPQPPATLQARDLPVYPALSLDSTD 58
 QY 61 PRAWLPTLQGLCTARAAQYLLHSPGLGASEAAAPRDEVDGRLVRRSSGAGSSXSPG 120
 Db 59 PRAWLPTLQGLCTARAAQYLLHSPGLGASEAAAPRDEADSQGLVLR--SCGGLSKSPG 117
 QY 121 PVKVRQLCKLKGVVVDELGCSRQAPSSKQVNGVQKRRRLAANARERRMHGHNHAFD 180
 Db 118 PVKVRQLCKLKGVVVDELGCSRQAPSSKQVNGVQKRRRLAANARERRMHGHNHAFD 177
 QY 181 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHHLRT 240
 Db 178 QLRNVIPSFNNDDKLSKYETLQMAQIYINALSELLQTPNVGEQPPPTASCKNDHHHLRT 237
 QY 241 AASVEGGAGNATAAGAAQAGSGSQRPPTPGSCRTFRFSAPASAGYSVQLDALHFSTFEDS 300
 Db 238 ASSVEGGAGASAVAGAQPAPGGGPRPTPGPCRTFRFSGASSGGYSVQLDALHFPFEDR 297
 QY 301 ALTAMMAQKNLSPSLPGSILOPVOEENSKTSPRSHRSDGFSFPHSHYSDSDEAS 354
 Db 298 ALTAMMAQKDLSPSLPGGILQPVQEDNSKTSPRSHRSDGFSFPHSHYSDSDEAS 351

RESULT 6
 AAB60369
 ID AAB60369 standard; protein; 351 AA.
 XX
 AC AAB60369;
 DT 24-APR-2001 (first entry)
 XX
 DE Mouse atonal homologue 1 (Math-1) protein, SEQ ID NO:46.
 XX
 KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.
 XX
 OS Mus musculus.
 XX
 PN WO200073764-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US015410.
 XX
 PR 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 XX
 DR WPI; 2001-032190/04.
 DR N-PSDB; AAF27277.
 XX
 PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or

```
CC amino acid sequence, or any of its homologues or orthologues as
CC therapeutic agents for the treatment of deafness, partial hearing loss,
CC vestibular effects due to damage or loss of inner hair cells,
CC osteoarthritis and abnormal cell proliferation. The invention also
CC encompasses methods of screening for compounds which affect the
CC expression of an atonal-associated nucleic acid sequence in an animal,
CC and a transgenic animal in which an allele of a native atonal-associated
CC gene is replaced by a heterologous nucleic acid sequence, thus
CC inactivating the atonal-associated allele. The nucleic acids or proteins
CC may be used in a method of treating an animal for hearing impairment,
CC joint disease, balance disorders, abnormal cell proliferation, or other
CC disease related to loss of a functional atonal-associated nucleic acid or
CC protein. They may particularly be used to treat an animal with a
CC deficiency in cerebellar granule neurons or their precursors, and may
CC also be used in promoting mechanoreceptive cell growth and generating
CC hair cells. The present sequence represents an atonal-associated amino
CC acid sequence referred to in the invention. Note: The present sequence is
CC not shown in the specification, but was obtained from GenBank
XX
SQ Sequence 351 AA;
Query Match 87.4%; Score 1627.5; DB 4; Length 351;
Best Local Similarity 88.4%; Pred. No. 5.4e-131;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
QY 1 MSRLHAEEWAEVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSLSDTD 60
DB 1 MSRLHAEEWAEVKELGDHHRQPQPHV--PPLTPQPPATLQARDLPVYPPELSLSDTD 58
QY 61 PRAWLTPTLQGLCTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGGASSKSPG 120
DB 59 PRAWLTPTLQGLCTARAAQYLLHSPELGASAAAAPRDEADSOGEVLR--SGCGGLSKSPG 117
QY 121 PVKVEQLCKLKGVVVDELGCSRQAPSSKQVGVQKORRLAANARERRRMHGLNHAFD 180
DB 118 PVKVEQLCKLKGVVVDELGCSRQAPSSKQVGVQKORRLAANARERRRMHGLNHAFD 177
QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELQTPSGGGQPPPPPPASCKSDHHHLRT 240
DB 178 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELQTPNVGQPPPTTASCKNDHHHLRT 237
QY 241 AASYEGGAGNATAAGAAQQAQSGGSGRTPPGSCRTFSAPASAGGYSVQLDALHFSFEDS 300
DB 238 ASSYEGGAGASAVAGAQAQAPGGGPRTPPGCTRFSPGASSGGYSVQLDALHFPFEDR 297
QY 301 ALTAMMAQKNLSPLSGSLQPVQEENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354
DB 298 ALTAMMAQKDLSPSLPGILQPVQEDNSKTSPRSHRSDGEFSPHSHYSDSDEAS 351
RESULT 7
AAO30939
ID AAO30939 standard; protein; 351 AA.
XX
AC AAO30939;
XX
DT 22-SEP-2003 (first entry)
XX
DE Mouse atonal homologue 1 (Math1) protein #13.
XX
KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
KW Math1; mouse.
XX
OS Mus musculus.
XX
PN WO2003047532-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US041458.
XX
PR 05-DEC-2001; 2001US-00004717.
XX
FA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Zoghbi HY, Qi Y;
XX
DR WPI; 2003-505253/47.
DR N-PSDB; AAL62224.
XX
PT New compositions comprising at least one stem cell which is up- or down-
PT regulated for expression of an atonal-associated sequence, and at least
PT one regulatory factor, useful for treating a gastrointestinal condition,
PT e.g. cancer.
XX
PS Disclosure; Page 144-145; 157pp; English.
XX
CC The invention relates to pharmaceutical compositions and methods for the
CC therapeutic use of an atonal-associated sequence. The composition
CC comprises of a stem cell which is up- or down-regulated for expression of
CC an atonal-associated sequence and a regulatory factor. Atonal-associated
CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
CC condition including cancer, damaged intestinal tissue, inflammatory bowel
CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
CC They are also useful for promoting mechanoreceptive cell growth, for
CC generating hair cells, for treating hearing impairment or imbalance
CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
CC disease that is a result or loss of functional atonal-associated
CC sequences. Atonal-associated sequences are also used in gene therapy. The
CC present sequence is mouse atonal homologue 1 (Math1) protein. This
CC sequence is used to illustrate the method of the invention
XX
SQ Sequence 351 AA;
Query Match 87.4%; Score 1627.5; DB 6; Length 351;
Best Local Similarity 88.4%; Pred. No. 5.4e-131;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
QY 1 MSRLHAEEWAEVKELGDHHRQPQPHHLPQPPPPQPPATLQAREHPVYPPELSLSDTD 60
DB 1 MSRLHAEEWAEVKELGDHHRQPQPHV--PPLTPQPPATLQARDLPVYPPELSLSDTD 58
QY 61 PRAWLTPTLQGLCTARAAQYLLHSPELGASAAAAPRDEVDGRGELVRRSSGGASSKSPG 120
DB 59 PRAWLTPTLQGLCTARAAQYLLHSPELGASAAAAPRDEADSOGEVLR--SGCGGLSKSPG 117
QY 121 PVKVEQLCKLKGVVVDELGCSRQAPSSKQVGVQKORRLAANARERRRMHGLNHAFD 180
DB 118 PVKVEQLCKLKGVVVDELGCSRQAPSSKQVGVQKORRLAANARERRRMHGLNHAFD 177
QY 181 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELQTPSGGGQPPPPPPASCKSDHHHLRT 240
DB 178 QLRNVIPSFNNKKLSKYETLQMAQIYINALSELQTPNVGQPPPTTASCKNDHHHLRT 237
QY 241 AASYEGGAGNATAAGAAQQAQSGGSGRTPPGSCRTFSAPASAGGYSVQLDALHFSFEDS 300
DB 238 ASSYEGGAGASAVAGAQAQAPGGGPRTPPGCTRFSPGASSGGYSVQLDALHFPFEDR 297
QY 301 ALTAMMAQKNLSPLSGSLQPVQEENSKTSPRSHRSDGEFSPHSHYSDSDEAS 354
DB 298 ALTAMMAQKDLSPSLPGILQPVQEDNSKTSPRSHRSDGEFSPHSHYSDSDEAS 351
RESULT 8
AAO30924
ID AAO30924 standard; protein; 351 AA.
XX
AC AAO30924;
XX
DT 22-SEP-2003 (first entry)
XX
DE Mouse atonal homologue 1 (Math1) protein #4.
XX
```

KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
 KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
 KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
 KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
 KW Math1; mouse.
 XX Mus musculus.
 OS
 XX WO2003047532-A2.
 FN
 XX 12-JUN-2003.
 PD
 XX
 XX 03-DEC-2002; 2002WO-US041458.
 PF
 XX 05-DEC-2001; 2001US-00004717.
 PR
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HY, Qi Y;
 PI WPI; 2003-505253/47.
 DR N-PSDB; AAL62204.
 DR
 XX New compositions comprising at least one stem cell which is up- or down-
 PT regulated for expression of an atonal-associated sequence, and at least
 PT one regulatory factor, useful for treating a gastrointestinal condition,
 PT e.g. cancer.
 XX
 XX Disclosure; Page 122-123; 157pp; English.
 XX
 CC The invention relates to pharmaceutical compositions and methods for the
 CC therapeutic use of an atonal-associated sequence. The composition
 CC comprises of a stem cell which is up- or down-regulated for expression of
 CC an atonal-associated sequence and a regulatory factor. Atonal-associated
 CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
 CC condition including cancer, damaged intestinal tissue, inflammatory bowel
 CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
 CC They are also useful for promoting mechanoreceptive cell growth, for
 CC generating hair cells, for treating hearing impairment or imbalance
 CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a
 CC disease that is a result of loss of functional atonal-associated
 CC sequences. Atonal-associated sequences are also used in gene therapy. The
 CC present sequence is mouse atonal homologue 1 (Math1) protein. This
 CC sequence is used to illustrate the method of the invention
 XX
 SQ Sequence 351 AA;
 Query Match 87.4%; Score 1627.5; DB 6; Length 351;
 Best Local Similarity 88.4%; Pred. No. 5.4e-131;
 Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;
 QY 1 MSRLHAEWAWEVKELGDRHQPHLPQPPPPPPPPATLQAREHPVVPPELSLSDTD 60
 DB 1 MSRLHAEWAWEVKELGDRHQPHV--PPLTPPPATLQARDLPVPAELSLSDTD 58
 QY 61 PRAWLPTLQICITARAQVLLHSPDELGASEAAAPRDEVDGRLVRRSGGASSKSPG 120
 DB 59 PRAWLPTLQGLCTARAQVLLHSPDELGASEAAAPRDEADSQGELVRR-SGCGGLSKSPG 117
 QY 121 PVKVRQLCKLGGVVVDLGGCRQAPSKQVNGKORRLAANARERRMHGLNHPD 180
 DB 118 PVKVRQLCKLGGVVVDLGGCRQAPSKQVNGKORRLAANARERRMHGLNHPD 177
 QY 181 QLRNVIPSNNDKKLSKYETLQMAQIYINALSELLQTPSGGQPPPPPPASCKSDHHLRT 240
 DB 178 QLRNVIPSNNDKKLSKYETLQMAQIYINALSELLQTPNVGQPPPTASCKNDHHLRT 237
 QY 241 AASYEGGAGNATAGQAQSGGQRTTPGSCRTFSPASAGGYVQLDALHFTFDS 300
 DB 238 AASYEGGAGASAVAGAPGPGGFRTPPGPCRTFRSGPSSGGYVQLDALHFPFEDR 297
 QY 301 ALTAMMAQKDLSPSLPGILQVQEDNSKTSRSHRSDGEFSPHSHYSDSEAS 354

Db 298 ALTAMMAQKDLSPSLPGILQVQEDNSKTSRSHRSDGEFSPHSHYSDSEAS 351
 RESULT 9
 AAB60366
 ID AAB60366 standard; protein; 266 AA.
 XX
 AC AAB60366;
 XX
 XX 24-APR-2001 (first entry)
 DT
 XX Zebrafish atonal homologue-1 (zath-1) protein, SEQ ID NO:40.
 DE
 XX Atonal; homologue; orthologue; atonal-associated protein; deafness;
 KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
 KW cellular proliferation; cerebellar granule neuron; gene therapy;
 KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
 KW transgenic animal.
 XX
 OS Danio rerio.
 XX
 XX WO2000073764-A2.
 FN
 XX 07-DEC-2000.
 PD
 XX 01-JUN-2000; 2000WO-US015410.
 PF
 XX 01-JUN-1999; 99US-0137060P.
 PR 19-JAN-2000; 2000US-0176993P.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 PA
 XX Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
 PI WPI; 2001-032190/04.
 DR N-PSDB; AAF27274.
 DR
 XX Therapeutic use of atonal-associated nucleic acids or amino acids, or any
 PT of its homologs or orthologs, for the treatment of e.g. deafness,
 PT osteoarthritis and abnormal cell proliferation.
 XX
 PS Disclosure; Page; 142pp; English.
 XX
 CC The invention relates to the use of atonal-associated nucleic acid or
 CC amino acid sequence, or any of its homologues or orthologues as
 CC therapeutic agents for the treatment of deafness, partial hearing loss,
 CC vestibular effects due to damage or loss of inner hair cells,
 CC osteoarthritis and abnormal cell proliferation. The invention also
 CC encompasses methods of screening for compounds which affect the
 CC expression of an atonal-associated nucleic acid sequence in an animal,
 CC and a transgenic animal in which an allele of a native atonal-associated
 CC gene is replaced by a heterologous nucleic acid sequence, thus
 CC inactivating the atonal-associated allele. The nucleic acids or proteins
 CC may be used in a method of treating an animal for hearing impairment,
 CC joint disease, balance disorders, abnormal cell proliferation, or other
 CC disease related to loss of a functional atonal-associated nucleic acid or
 CC protein. They may particularly be used to treat an animal with a
 CC deficiency in cerebellar granule neurons or their precursors, and may
 CC also be used in promoting mechanoreceptive cell growth and generating
 CC hair cells. The present sequence represents an atonal-associated amino
 CC acid sequence referred to in the invention. Note: The present sequence is
 CC not shown in the specification, but was obtained from GenBank
 XX
 SQ Sequence 266 AA;
 Query Match 31.3%; Score 582; DB 4; Length 266;
 Best Local Similarity 46.5%; Pred. No. 1.8e-41;
 Matches 155; Conservative 21; Mismatches 69; Indels 88; Gaps 14;
 QY 12 EYKELGDHHRQPHLPQPPPPPPATLQAREHPVVPPELSLSDTDPRAWLPTLQ 71
 DB 10 EVELLDVQH-----SSLRGESQKYPALMALMASSDPRAWLAPVQAG 51

CC hair cells. The present sequence represents an atonal-associated amino
CC acid sequence referred to in the invention. Note: The present sequence is
CC not shown in the specification, but was obtained from GenBank

XX
SQ Sequence 161 AA;

Query Match 18.1%; Score 337; DB 4; Length 161;
Best Local Similarity 61.2%; Pred. No. 9.3e-21;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSGGAS----SKSPGVKVRQCLKLGKGVVDELGCSRQAPSSKQVGVKQRRLLA 163
Db 42 RPVAGGARVPAGAPGE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRRLLA 89

QY 164 ANARERRMGLNHFADQLRNVPFNNNDKLSKYETLQMAQIYINALSSELLQTPPSGGEQ 223
Db 90 ANARERRMGLNHFADQLRNVPFNNNDKLSKYETLQMAQIYISALAEHLHGPPA--- 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAXAE 155

RESULT 13
AAO30946
ID AAO30946 standard; protein; 161 AA.
XX
AC AAO30946;
XX
DT 22-SEP-2003 (first entry)
XX
DE Chicken atonal homologue 1 (Cathl) protein #4.
XX
KW Atonal-associated protein; gastrointestinal; damaged intestinal tissue;
KW cancer; inflammatory bowel disease; irritable bowel syndrome; infection;
KW necrotising enterocolitis; hearing impairment; abnormal proliferation;
KW joint disease; gene therapy; imbalance disorder; atonal homologue 1;
KW Cathl; chicken.
XX
OS Gallus sp.
XX
PN WO2003047532-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-US041459.
XX
PR 05-DEC-2001; 2001US-00004717.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Zoghbi HY, Qi Y;
XX
DR WPI; 2003-505253/47.
DR N-PSDB; AAL62231.
XX

XX New compositions comprising at least one stem cell which is up- or down-
PT regulated for expression of an atonal-associated sequence, and at least
PT one regulatory factor, useful for treating a gastrointestinal condition,
PT e.g. cancer.

XX
PS Disclosure; Page 152-153; 157pp; English.
XX
CC The invention relates to pharmaceutical compositions and methods for the
CC therapeutic use of an atonal-associated sequence. The composition
CC comprises of a stem cell which is up- or down-regulated for expression of
CC an atonal-associated sequence and a regulatory factor. Atonal-associated
CC amino acid or nucleic acid sequence is used to treat a gastrointestinal
CC condition including cancer, damaged intestinal tissue, inflammatory bowel
CC disease, irritable bowel syndrome, infection or necrotising enterocolitis.
CC They are also useful for promoting mechanoreceptive cell growth, for
CC generating hair cells, for treating hearing impairment or imbalance
CC disorder, joint disease, abnormal proliferation (e.g. cancer), or a

CC disease that is a result or loss of functional atonal-associated
CC sequences. Atonal-associated sequences are also used in gene therapy. The
CC present sequence is chicken atonal homologue 1 (Cathl) protein. This
CC sequence is used to illustrate the method of the invention

XX
SQ Sequence 161 AA;

Query Match 18.1%; Score 337; DB 6; Length 161;
Best Local Similarity 61.2%; Pred. No. 9.3e-21;
Matches 79; Conservative 10; Mismatches 22; Indels 18; Gaps 5;

QY 107 RRSGGAS----SKSPGVKVRQCLKLGKGVVDELGCSRQAPSSKQVGVKQRRLLA 163
Db 42 RPVAGGARVPAGAPGE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRRLLA 89

QY 164 ANARERRMGLNHFADQLRNVPFNNNDKLSKYETLQMAQIYINALSSELLQTPPSGGEQ 223
Db 90 ANARERRMGLNHFADQLRNVPFNNNDKLSKYETLQMAQIYISALAEHLHGPPA--- 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAXAE 155

RESULT 14
AAB60356
ID AAB60356 standard; protein; 178 AA.
XX
AC AAB60356;
XX
DT 24-APR-2001 (first entry)
XX
DE Chicken atonal homologue ngn1/ath4c protein, SEQ ID NO:19.
XX
KW Atonal; homologue; orthologue; atonal-associated protein; deafness;
KW hearing impairment; vestibular effect; balance disorder; osteoarthritis;
KW cellular proliferation; cerebellar granule neuron; gene therapy;
KW mechanoreceptive cell growth; auditory; osteopathic; cytostatic;
KW transgenic animal.
XX
OS Gallus gallus.
XX
PN WO2000073764-A2.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US015410.
XX
PR 01-JUN-1999; 99US-0137060P.
PR 19-JAN-2000; 2000US-0176993P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX
PI Zoghbi HY, Bellen H, Birmingham N, Hassan B, Ben-Arie N;
XX
DR WPI; 2001-032190/04.
DR N-PSDB; AAF27263.
XX
PT Therapeutic use of atonal-associated nucleic acids or amino acids, or any
PT of its homologs or orthologs, for the treatment of e.g. deafness,
PT osteoarthritis and abnormal cell proliferation.
XX
PS Disclosure; Page; 142pp; English.

XX
CC The invention relates to the use of atonal-associated nucleic acid or
CC amino acid sequence, or any of its homologues or orthologues as
CC therapeutic agents for the treatment of deafness, partial hearing loss,
CC vestibular effects due to damage or loss of inner hair cells,
CC osteoarthritis and abnormal cell proliferation. The invention also
CC encompasses methods of screening for compounds which affect the
CC expression of an atonal-associated nucleic acid sequence in an animal,
CC and a transgenic animal in which an allele of a native atonal-associated
CC gene is replaced by a heterologous nucleic acid sequence, thus

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:18:48 ; Search time 43.424 Seconds
(without alignments)
420.864 Million cell updates/sec

Title: US-09-980-381A-58

Perfect score: 1862
Sequence: 1 MSRLHAHAEWAWEKELGDHH.....HRSDGEPSPHSHYSDSDEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	15.1	57	4	US-08-722-570-8
2	281	15.1	57	4	US-08-932-411A-8
3	217	11.7	263	4	US-08-932-411A-18
4	214	11.5	352	1	US-08-552-142A-4
5	214	11.5	352	1	US-08-910-973-4
6	214	11.5	352	4	US-09-499-227-4
7	214	11.5	352	5	PCT-US95-05741-4
8	210.5	11.3	381	1	US-08-910-973-11
9	210.5	11.3	381	4	US-09-499-227-11
10	210	11.3	382	3	US-09-234-332-9
11	209.5	11.3	237	1	US-08-910-973-13
12	209.5	11.3	237	4	US-09-499-227-13
13	206.5	11.1	379	1	US-08-552-142A-11
14	205	11.0	244	4	US-08-722-570-1
15	205	11.0	244	4	US-08-932-411A-1
16	204	11.0	380	1	US-08-552-142A-17
17	202	10.8	244	4	US-08-910-973-22
18	202	10.8	244	4	US-09-499-227-22
19	202	10.8	244	4	US-08-932-411A-14
20	202	10.8	357	1	US-08-552-142A-2
21	202	10.8	357	1	US-08-910-973-2
22	202	10.8	357	4	US-09-499-227-2
23	202	10.8	357	5	PCT-US95-05741-2
24	201.5	10.8	214	4	US-08-932-411A-20
25	201.5	10.8	382	1	US-08-910-973-17
26	201.5	10.8	382	4	US-09-499-227-17
27	200	10.7	356	1	US-08-552-142A-15

28	200	10.7	356	1	US-08-910-973-15	Sequence 15, Appl
29	200	10.7	356	4	US-09-499-227-15	Sequence 15, Appl
30	199	10.7	356	3	US-09-234-332-7	Sequence 7, Appl
31	196.5	10.6	214	4	US-08-722-570-2	Sequence 2, Appl
32	196.5	10.6	214	4	US-08-932-411A-2	Sequence 2, Appl
33	196	10.5	57	4	US-08-722-570-9	Sequence 9, Appl
34	196	10.5	57	4	US-08-932-411A-9	Sequence 9, Appl
35	194	10.4	356	3	US-09-234-332-8	Sequence 8, Appl
36	182	9.8	156	1	US-08-552-142A-9	Sequence 9, Appl
37	182	9.8	156	1	US-08-910-973-9	Sequence 9, Appl
38	182	9.8	156	4	US-09-499-227-9	Sequence 9, Appl
39	182	9.8	156	5	PCT-US95-05741-9	Sequence 9, Appl
40	173	9.3	57	4	US-08-722-570-5	Sequence 5, Appl
41	173	9.3	57	4	US-08-932-411A-5	Sequence 5, Appl
42	169	9.1	57	4	US-08-722-570-6	Sequence 6, Appl
43	169	9.1	57	4	US-08-932-411A-6	Sequence 6, Appl
44	166	8.9	161	5	PCT-US95-05741-11	Sequence 11, Appl
45	162	8.7	57	4	US-08-722-570-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-722-570-8
; Sequence 8, Application US/08722570
; Patent No. 6555337
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-722-570-8

Query Match 15.1%; Score 281; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2.5e-19;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 158 KORRLAANARERRRMHGLNHFQDLNRNVPFNNDKLSKYETLQMAQIYINALSEL 214

Db 1 KNRLAANARERRRMHGLNHFQDLNRNVPFNNDKLSKYETLQMAQIYINALSEI 57

```

;
; TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
;
; TITLE OF INVENTION: and Proteins
;
; NUMBER OF SEQUENCES: 20

```


us-09-980-381a-58.ra1

Wed Sep 22 12:21:30 2004

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/499,227
 FILING DATE: 05-August-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/239,238
 FILING DATE: 06-May-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US95/05741
 FILING DATE: 08-May-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/17532
 FILING DATE: 30-October-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/910,973
 FILING DATE: 07-August-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Sheiness, Diana K.
 REGISTRATION NUMBER: 35,356
 REFERENCE/DOCKET NUMBER: FHCR-1-12742
 TELEPHONE: 206-682-8100; 206-224-0735 (direct)
 TELEFAX: 206-225-0779
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 381 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-09-499-227-11

Query Match 11.3%; Score 210.5; DB 4; Length 381;
 Best Local Similarity 26.4%; Pred. No. 1.7e-11;
 Matches 98; Conservative 42; Mismatches 126; Indels 105; Gaps 17;

QY	22	QPOPHLPQPPPPPPPPATLQAREHVPVPELSLLDTPRAWLAPTLOGICTARAAQYL	81
DB	28	EPKSGKGDAPPPPPAP-----GPGAPG--PARAAKPV	58
QY	82	LHSPELGASAAAPRDEVDGRLVRRSSGGASSKSPGVKVRQELCKLGGVVVDL	141
DB	59	---PLRGEETATLAEVKEGEL-----GEEEEEE-----EEEGLEAEG-----E	99
QY	142	CSRQAPSSKQVGVQ-----KORLAANARERRMHGLNHAFLDQLRNVIPSFNNDKLSK	197
DB	100	RPKRGPKKRMTKARLERSKLKQKANARERNRMDLNAALDNLKRVVPCYKTKQLSK	159
QY	198	YETLQMAQIYNALSELL-----QTPSGGEQPPPP--PASC-KSDHHLRTAA	242
DB	160	IETLRLAKNYIWLSEILRSKRPDLVSVVQTLCKGLSQPTNLVAGCLQNSRNFLEQ	219
QY	243	SYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPGSCRTFSAPA	280
DB	220	GADGAGRPHGSGGPFAMHPYPYPCSLAGACQAAAGLGGGAHAALRTHGYCAAYETLYA	279
QY	281	SAGYSVQLDALHFTFEDSALTAMMAQNLSPSLPGSILQPVQENSKTSPRSHRSDGE	340
DB	280	AAGGGASPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPDHEKSVHYSMH	328
QY	341	FSP-----HSH	346
DB	329	YSALPGSRHG	339

RESULT 10
 US-09-234-332-9
 ; Sequence 9, Application US/09234332A
 ; Patent No. 6087168
 ; GENERAL INFORMATION:

APPLICATION NUMBER: WO PCT/US95/05741
 FILING DATE: 08-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/17532
 FILING DATE: 30-October-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Sheiness, Diana K.
 REGISTRATION NUMBER: 35,356
 REFERENCE/DOCKET NUMBER: FHCR-1-10958
 TELEPHONE: 206-682-8100; 206-224-0735 (direct)
 TELEFAX: 206-225-0779
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 381 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-910-973-11

Query Match 11.3%; Score 210.5; DB 1; Length 381;
 Best Local Similarity 26.4%; Pred. No. 1.7e-11;
 Matches 98; Conservative 42; Mismatches 126; Indels 105; Gaps 17;

QY	22	QPOPHLPQPPPPPPPPATLQAREHVPVPELSLLDTPRAWLAPTLOGICTARAAQYL	81
DB	28	EPKSGKGDAPPPPPAP-----GPGAPG--PARAAKPV	58
QY	82	LHSPELGASAAAPRDEVDGRLVRRSSGGASSKSPGVKVRQELCKLGGVVVDL	141
DB	59	---PLRGEETATLAEVKEGEL-----GEEEEEE-----EEEGLEAEG-----E	99
QY	142	CSRQAPSSKQVGVQ-----KORLAANARERRMHGLNHAFLDQLRNVIPSFNNDKLSK	197
DB	100	RPKRGPKKRMTKARLERSKLKQKANARERNRMDLNAALDNLKRVVPCYKTKQLSK	159
QY	198	YETLQMAQIYNALSELL-----QTPSGGEQPPPP--PASC-KSDHHLRTAA	242
DB	160	IETLRLAKNYIWLSEILRSKRPDLVSVVQTLCKGLSQPTNLVAGCLQNSRNFLEQ	219
QY	243	SYEG-----GAGNATA-----AGAQ-QASGG-----SORPTPGSCRTFSAPA	280
DB	220	GADGAGRPHGSGGPFAMHPYPYPCSLAGACQAAAGLGGGAHAALRTHGYCAAYETLYA	279
QY	281	SAGYSVQLDALHFTFEDSALTAMMAQNLSPSLPGSILQPVQENSKTSPRSHRSDGE	340
DB	280	AAGGGASPD--YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPDHEKSVHYSMH	328
QY	341	FSP-----HSH	346
DB	329	YSALPGSRHG	339

RESULT 9
 US-09-499-227-11
 ; Application US/09499227
 ; Patent No. 6444463
 ; GENERAL INFORMATION:
 ; APPLICANT: Tapscott, Stephen J.
 ; APPLICANT: Olson, James M.
 ; TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
 ; STREET: 1420 Fifth Avenue, Suite 2800
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98101-2347
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy

```
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; OTHER INFORMATION: Neurogenic basic helix-loop-helix protein (neuro
; OTHER INFORMATION: D2); Genbank Accession U58681.
US-09-234-332-9

Query Match      11.3%; Score 210; DB 3; Length 382;
Best Local Similarity 26.6%; Pred. No. 1.9e-11;
Matches 98; Conservative 41; Mismatches 119; Indels 110; Gaps 17;

QY 22 QQPHLPPPPPPPPATLQAREHPVPPPELSDSTDPRAWLPTLOGICTARAAQYL 81
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
28 EPRSDKGDAAPPPPPPPAP-----GPGAPG--PARAAKPV 58

QY 82 LHSPELCASAAAPRDEVDGRLVRRSSGASSKSPGVKREOLCKLKGVVVDELG 141
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 ---PLRGEGTETALAEVKEGEL-----GGEEREE-----E 99

QY 142 CSRQAPSSKQNGVQ----KORRLAANARERRMGLNHAFDQLRNVPSPFNNDKLSK 197
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
100 RPKRGPKRKMTKARLERSKLRRQCANARERNMHDNALDNLRKVPCKYKQKLSK 159

QY 198 YETLQWAIYINALSELL-----QTPSGGEQPPPP--PASC-KSDHHLRTAA 242
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
160 IETLAKNYIWALESILLRSGKRPDLVSYVOTLCKGLSQPTNLVAGCLQLNSRNFLTEQ 219

QY 243 SYEG-----GAGNATA-----AGAQ-OASGG-----SQRTPPGSCRTFSAPA 280
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
220 GADGAGRHGSGGPPFAMHPYPYPCSLAGAQQAAGLGGGAHALRTHGYCAAYETLYA 279

QY 281 SAGYSVOLDALHFTFDSALTAMMAQKNLSPSLFGSILQPVQENSKTSRSHRSDGE 340
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
280 AAGGGGASPD-YNSSEYEGPLSPPLCLNGNFS-----LKQDSSPDHEKSY----- 323

QY 341 FSPSHYS 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
324 -----HYS 326

RESULT 11
US-09-910-973-13
; Sequence 13, Application US/08910973
; Patent No. 5795723
; GENERAL INFORMATION:
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Olson, James M.
; TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; OTHER INFORMATION: Neurogenic basic helix-loop-helix protein (neuro
; OTHER INFORMATION: D2); Genbank Accession U58681.
US-09-234-332-9

Query Match      11.3%; Score 209.5; DB 1; Length 237;
Best Local Similarity 33.8%; Pred. No. 1.1e-11;
Matches 70; Conservative 26; Mismatches 64; Indels 47; Gaps 9;

QY 107 RRSSGGAS--SSKSPGVKREOLCKLKGVVVDELGCSRQAPSSKQNGVQKORRLAAN 165
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 RRGAPNISRASEVPGAQDDEQERRRRG-----RTRVRSALLHSLSRRSRVKAN 98

QY 166 ARERERMGLNHAFDQLRNVPSPFNNDKLSKYETLQWAIYINALSELLQ-----TPSG 220
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
99 DRERNMNLNAALDALRSVLSPFPDPTKLTKIETLRFAYNIWALAETRLADQGLPGG 158

QY 221 GEQ-----PP-----PPASCKSDHHLRTAAASYEGGAGNATAAGAAQASGSGQ----RP 266
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
159 GARELLPQCVPCLPGPPSPASD-----AESWGSAGAAASPLSDPSPAASEDTYRP 212

QY 267 TTPGSCRTFSAPAASAGYSVOLDALH 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 GDP-----VFSFP-----SLPKDLIH 228

RESULT 12
US-09-499-227-13
; Sequence 13, Application US/09499227
; Patent No. 644463
; GENERAL INFORMATION:
; APPLICANT: Tapscott, Stephen J.
; APPLICANT: Olson, James M.
; TITLE OF INVENTION: Expression of Neurogenic bHLH Genes in Primitive Neuroectoder
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson KindnessPLLC
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```


Wed Sep 22 12:21:30 2004

APPLICATION NUMBER: US/09/499,227
FILING DATE: 05-August-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17532
FILING DATE: 30-October-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/910,973
FILING DATE: 07-August-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHC-1-12742
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100; 206-224-0735 (direct)
TELEFAX: 206-225-0779
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-499-227-13

Query Match 11.3%; Score 209.5; DB 4; Length 237;
Best Local Similarity 33.8%; Pred. No. 1.1e-11;
Matches 70; Conservative 26; Mismatches 64; Indels 47; Gaps 9;
QY 107 RASGGAS-SKSGPGVKVREQLCKLKGVVVDELGCSQRAPSKQVNGVQKQRLAAN 165
DB 49 RGAENIASEVPGAQDEQERRRG-----RTRVRSALLHLRSRRVKAN 98
QY 166 ARERRMHGLNHFQDLNRNIPSPNNDKLSKYETLQMAQIYNALSELLQ-----TPSG 220
DB 99 DERNRMNNAALDARSVLSPFPDDTKLTETLRFAYNIWALAEFLRLADQGLGG 158
QY 221 GQO-----PPASCKSDHHHLRTAASVGGAGNATAAGAAQASGGSQ-----RP 266
DB 159 GARERLLPQCVPCLFGPPSPASD-----AESWGSAAAASPLDPPSPAASEFTYRP 212
QY 267 TPGSCRTFSAPASAGGYSVOLDALH 293
DB 213 GDP-----VFSPP-----SLPKDLH 228

RESULT 13
US-08-552-142A-11
Sequence 11, Application US/08552142A
Patent No. 569595
GENERAL INFORMATION:
APPLICANT: Weintraub, Harold M.
APPLICANT: Lee, Jacqueline E.
APPLICANT: Tapscott, Stephen J.
APPLICANT: Hollenberg, Stanley M.
TITLE OF INVENTION: Neurogenic Differentiation (NeuroD) Genes
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/552,142A
FILING DATE: 02-NOV-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/239,238
FILING DATE: 06-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US95/05741
FILING DATE: 08-May-1995
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: FHC-1-8933
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-225-0709
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-552-142A-11

Query Match 11.1%; Score 206.5; DB 1; Length 379;
Best Local Similarity 25.5%; Pred. No. 4e-11;
Matches 94; Conservative 38; Mismatches 134; Indels 103; Gaps 14;
QY 22 QPQHLLPQPPPPPPATLQAREHPVYPPELSLIDTDPRAWLPTLQGGICTARAAQYL 81
DB 28 EPRSDKGADPPPPPPAP-----GPGAPG--PARAAKPV 58
QY 82 LHSPELGASEAAAPRDEVDGELVRRSSGASSKSPGVKVRQLCKLKGVVVDELG 141
DB 59 ---PLRGEETATLAEVKEGEL---GGESEEE---EEERGLDEAG-----E 99
QY 142 CSRQAPSKQVNGVQ---KORLAANARERRRMHGLNHFQDLNRNIPSPNNDKLSK 197
DB 100 RPKRGPKRKMTKARLERSKLRQKANARERNRMDLNAALDNRKVVPCYSKTKQLSK 159
QY 198 YETLQMAQIYNALSELL-----QTPSGEQPPPP--PASC-KSDHHHLRTAA 242
DB 160 IETRLAKNIWALSEILRSGRPDLSVYVQTLCKGLSOPTTNLVAGCLQLNSRNFLTEQ 219
QY 243 SYEGGA-----GNATAAGAAQASGGSORPTTPGSCRTFSAPASA 282
DB 220 GRDGAXRFHSGGPPFAMHPYPYPCSRGGRTVFGAAAWAAGARLRTGYYCAAYETLYAA 279
QY 283 GGYSVOLDALHFTSTFSDSALTAMMAQKNLSPSLPGSLQPOVEENSKTSPRS-----HRS 337
DB 280 GGGGASPD--YNSSEYEGPLSPCLCLNGNFS-----LKQDSSPDHEKSHYHSMHYS 328
QY 338 DGEFSPSH 346
DB 329 GCPGSRHGH 337

RESULT 14
US-08-722-570-1
Sequence 1, Application US/08722570
Patent No. 655337
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Ma, Qifu
TITLE OF INVENTION: NEUROGENIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400

```

; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-722-570-1

Query Match      11.0%; Score 205; DB 4; Length 244;
Best Local Similarity 34.2%; Pred. No. 2.9e-11;
Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;

QY 95 PRDEVDRGELVRRSSGGAS-SSKSPGPVKVREQLCKLGGVVVDELGCSRQAPSSKQV 153
Db 38 PLASTSGLSVPARRSAPTLSGASNVPGQDEEGERRRRG-----RARVRSEALL 87
QY 154 NGVQKORRLAANARERRRMHGLNHFQDLNRNIPSFNNDKKLSKYETLQMAQIYINALSE 213
Db 88 HSLRRSRVKANDRRERNMNLNAALDALRSVLPSPDDTKLTKIETLRFAYNIWALAE 147
QY 214 LLQ-----TPSGEQ-----PP-----PPASCKSDHHHLRTAASVEGGAGNATAAGAOQ 258
Db 148 TURLADQGLPGGARERLLPQCVCPLPGPPSPASD-----TESWGSAAASPCATV 199
QY 259 ASGGSQRPTPPGS 271
Db 200 ASPLSDPSPSPAS 212

```

RESULT 15

```

US-08-932-411A-1
; Sequence 1, Application US/08932411A
; Patent No. 6566496
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Ma, Qifu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3-RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-932-411A-1

```

```

Query Match      11.0%; Score 205; DB 4; Length 244;
Best Local Similarity 34.2%; Pred. No. 2.9e-11;
Matches 66; Conservative 24; Mismatches 69; Indels 34; Gaps 6;

QY 95 PRDEVDRGELVRRSSGGAS-SSKSPGPVKVREQLCKLGGVVVDELGCSRQAPSSKQV 153
Db 38 PLASTSGLSVPARRSAPTLSGASNVPGQDEEGERRRRG-----RARVRSEALL 87
QY 154 NGVQKORRLAANARERRRMHGLNHFQDLNRNIPSFNNDKKLSKYETLQMAQIYINALSE 213
Db 88 HSLRRSRVKANDRRERNMNLNAALDALRSVLPSPDDTKLTKIETLRFAYNIWALAE 147
QY 214 LLQ-----TPSGEQ-----PP-----PPASCKSDHHHLRTAASVEGGAGNATAAGAOQ 258
Db 148 TURLADQGLPGGARERLLPQCVCPLPGPPSPASD-----TESWGSAAASPCATV 199
QY 259 ASGGSQRPTPPGS 271
Db 200 ASPLSDPSPSPAS 212

```

Search completed: September 21, 2004, 21:25:51
Job time : 44.424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 21:24:21 ; Search time 161.424 Seconds
(without alignments)
704.240 Million cell updates/sec

Title: US-09-980-381A-58

Perfect score: 1862
Sequence: 1 MSRLHAEAEVKEVKGDLHH.....HRSDGEFSPHSHYSDEAS 354

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1862	100.0	354	13	US-10-004-717-2
2	1862	100.0	354	13	US-10-004-717-58
3	1862	100.0	354	16	US-10-373-249-3
4	1862	100.0	354	16	US-10-373-249-4
5	1627.5	87.4	351	13	US-10-004-717-11
6	1627.5	87.4	351	13	US-10-004-717-46
7	1627.5	87.4	351	16	US-10-373-249-1
8	1627.5	87.4	351	16	US-10-373-249-2
9	582	31.3	266	13	US-10-004-717-40
10	537	28.8	259	13	US-10-004-717-66
11	337	18.1	161	13	US-10-004-717-60
12	281	15.1	57	8	US-08-722-570-8
13	281	15.1	57	12	US-10-425-259-8
14	228.5	12.3	178	13	US-10-004-717-19
15	224.5	12.1	213	13	US-10-004-717-21

16	222	11.9	250	16	US-10-642-093-25	Sequence 25, Appl
17	222	11.9	250	16	US-10-795-002-25	Sequence 25, Appl
18	219	11.8	208	12	US-10-403-571-12	Sequence 12, Appl
19	218	11.7	263	13	US-10-004-717-7	Sequence 7, Appl
20	218	11.7	263	13	US-10-004-717-38	Sequence 38, Appl
21	217	11.7	151	13	US-10-004-717-36	Sequence 36, Appl
22	217	11.7	263	12	US-10-425-259-18	Sequence 18, Appl
23	217	11.7	263	13	US-10-004-717-31	Sequence 31, Appl
24	214	11.5	352	16	US-10-654-102-6	Sequence 6, Appl
25	214	11.5	352	16	US-10-654-102-16	Sequence 16, Appl
26	214	11.5	352	16	US-10-654-102-27	Sequence 27, Appl
27	214	11.5	352	16	US-10-654-102-38	Sequence 38, Appl
28	214	11.5	352	16	US-10-654-102-43	Sequence 43, Appl
29	213.5	11.5	237	16	US-10-642-093-23	Sequence 23, Appl
30	213.5	11.5	237	16	US-10-654-102-42	Sequence 42, Appl
31	213.5	11.5	237	16	US-10-795-002-23	Sequence 23, Appl
32	213	11.4	312	13	US-10-004-717-64	Sequence 64, Appl
33	212.5	11.4	189	13	US-10-004-717-17	Sequence 17, Appl
34	211.5	11.4	325	15	US-10-120-801-96	Sequence 96, Appl
35	210.5	11.3	381	15	US-10-435-696-32	Sequence 32, Appl
36	210.5	11.3	381	16	US-10-654-102-19	Sequence 19, Appl
37	210	11.3	350	16	US-10-654-102-11	Sequence 11, Appl
38	210	11.3	350	16	US-10-654-102-12	Sequence 12, Appl
39	210	11.3	350	16	US-10-654-102-13	Sequence 13, Appl
40	210	11.3	382	16	US-10-654-102-31	Sequence 31, Appl
41	209.5	11.3	237	16	US-10-642-093-21	Sequence 21, Appl
42	209.5	11.3	237	16	US-10-795-002-21	Sequence 21, Appl
43	207.5	11.1	355	16	US-10-654-102-9	Sequence 9, Appl
44	207	11.1	138	13	US-10-004-717-50	Sequence 50, Appl
45	206.5	11.1	379	16	US-10-654-102-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-004-717-2
; Sequence 2, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-717-2

Query Match 100.0%; Score 1862; DB 13; Length 354;

Best Local Similarity 100.0%; Pred. No. 1.7e-135;

Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSRLHAEAEVKEVKGDLHHRQPHLPQPPPPQPPATLQAEHVPVPELSLLDSTD 60

Db 1 MSRLHAEAEVKEVKGDLHHRQPHLPQPPPPQPPATLQAEHVPVPELSLLDSTD 60

Qy 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAPRDEVDGRGELVRRSGGSSSSKSPG 120

Db 61 PRAWLAPTLOGICTARAAQYLLHSPELGASAAAPRDEVDGRGELVRRSGGSSSSKSPG 120

Qy	121	PVKYRBOI	CKLKGKGVVVD	ELGCSRORAPSSK	QVNGVQKQRRLA	ANARERRRMHGLN	HAED	180
Db	121	PVKYREQI	CKLKGKGVVVD	ELGCSRORAPSSK	QVNGVQKQRRLA	ANARERRRMHGLN	HAED	180
Qy	181	QLRNVIFSN	NDKKLSKYETI	LQMAQIYNALSEL	LQTPSGGQBP	PPPPASCCKSDHH	LRT	240
Db	181	QLRNVIFSN	NDKKLSKYETI	LQMAQIYNALSEL	LQTPSGGQBP	PPPPASCCKSDHH	LRT	240
Qy	241	AASVEGGAG	NATAAGAAQ	AGSGQRPTPGSC	RTRFSAPASAGG	YSVQLDALHFT	FEDES	300
Db	241	AASVEGGAG	NATAAGAAQ	AGSGQRPTPGSC	RTRFSAPASAGG	YSVQLDALHFT	FEDES	300
Qy	301	ALTANMAQ	KNLSPSIPG	SILPQVEENSKT	SPRSHRSDGEF	SPHSHYSDSD	DEAS	354
Db	301	ALTANMAQ	KNLSPSIPG	SILPQVEENSKT	SPRSHRSDGEF	SPHSHYSDSD	DEAS	354

QY 61 PRAWLAPTLOGICTARAAQYLLHSPDLGASAAAAPRDEVDGRELVRSSGASSSKSPG 120
Db 61 PRAWLAPTLOGICTARAAQYLLHSPDLGASAAAAPRDEVDGRELVRSSGASSSKSPG 120
QY 121 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 180
Db 121 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 180
QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
Db 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
QY 241 AASYEGGAGNATAAGAAQASGSGORPTPGSCRTFRSAPASAGGYSVOLDALHFSFTFDS 300
Db 241 AASYEGGAGNATAAGAAQASGSGORPTPGSCRTFRSAPASAGGYSVOLDALHFSFTFDS 300
QY 301 ALTAMMAOKNLSPLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354
Db 301 ALTAMMAOKNLSPLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354

RESULT 5

US-10-004-717-11

; Sequence 11, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHEI, HUDA Y.
; APPLICANT: YANG, QI

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 11
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus

US-10-004-717-11

Query Match 87.4%; Score 1627.5; DB 13; Length 351;
Best Local Similarity 88.4%; Pred. No. 2.1e-117;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEWEAEVKELGDHHRQPQPHLPQPPPPPPATLQAREHPVVPPELSLLDSTD 60
Db 1 MSRLHAEWEAEVKELGDHHRQPQPHV--PPLTQPPATLQARDLPVYPAELSLLDSTD 58
QY 61 PRAWLAPTLOGICTARAAQYLLHSPDLGASAAAAPRDEVDGRELVRSSGASSSKSPG 120
Db 59 PRAWLPTTLOGICTARAAQYLLHSPDLGASAAAAPRDEADSOQELVRR--SGCGGLSKSPG 117
QY 121 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 180
Db 118 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 177
QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
Db 178 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPNVGQPPPTASCKNDHHHLRT 237
QY 241 AASYEGGAGNATAAGAAQASGSGORPTPGSCRTFRSAPASAGGYSVOLDALHFSFTFDS 300
Db 238 ASSYEGGAGASAVAGAQAAPGCGGPRPTPGPCRTFRSGPSSGGYSVOLDALHFPAPEDR 297

QY 301 ALTAMMAOKNLSPLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354
Db 298 ALTAMMAOKNLSPLPGSILQPVQEDNKTSPRSHRSDGFEFSPHSHYSDEAS 351
RESULT 6
US-10-004-717-46
; Sequence 46, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHEI, HUDA Y.
; APPLICANT: YANG, QI

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19

; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 46
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Mus musculus

US-10-004-717-46

Query Match 87.4%; Score 1627.5; DB 13; Length 351;
Best Local Similarity 88.4%; Pred. No. 2.1e-117;
Matches 313; Conservative 11; Mismatches 27; Indels 3; Gaps 2;

QY 1 MSRLHAEWEAEVKELGDHHRQPQPHLPQPPPPPPATLQAREHPVVPPELSLLDSTD 60
Db 1 MSRLHAEWEAEVKELGDHHRQPQPHV--PPLTQPPATLQARDLPVYPAELSLLDSTD 58

QY 61 PRAWLAPTLOGICTARAAQYLLHSPDLGASAAAAPRDEVDGRELVRSSGASSSKSPG 120
Db 59 PRAWLPTTLOGICTARAAQYLLHSPDLGASAAAAPRDEADSOQELVRR--SGCGGLSKSPG 117

QY 121 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 180
Db 118 PVKREOLCKLGGVVDELGCSRORAPSSKQVNGVQKORRLAANARERRRMHGLNHAED 177

QY 181 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPSGGEGQPPPPASCKSDHHHLRT 240
Db 178 QLRNVIPSFNNDKKLSKYETLQMAQIYINALSELLQTPNVGQPPPTASCKNDHHHLRT 237

QY 241 AASYEGGAGNATAAGAAQASGSGORPTPGSCRTFRSAPASAGGYSVOLDALHFSFTFDS 300
Db 238 ASSYEGGAGASAVAGAQAAPGCGGPRPTPGPCRTFRSGPSSGGYSVOLDALHFPAPEDR 297

QY 301 ALTAMMAOKNLSPLPGSILQPVQENSKTSPRSHRSDGFEFSPHSHYSDEAS 354
Db 298 ALTAMMAOKNLSPLPGSILQPVQEDNKTSPRSHRSDGFEFSPHSHYSDEAS 351

RESULT 7

US-10-373-249-1

; Sequence 1, Application US/10373249
; Publication No. US20040166091A1
; GENERAL INFORMATION:

; APPLICANT: BROUGH, DOUGLAS E
; TITLE OF INVENTION: MATERIALS AND METHODS FOR TREATING DISORDERS OF THE EAR

; FILE REFERENCE: 214680
; CURRENT APPLICATION NUMBER: US/10/373,249
; CURRENT FILING DATE: 2003-02-24

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1

QY	241	AAAYEGGAGNATAAGAAQAGGSGQRTPPGSCRTFRFSAPASAGGYSVQDLALHFSFTFDS	300
Db	238	ASSYEGGAGASAVAGAPAGGGPRTPPGCCTRFGSGASSGYSVQDLALHFPAPEDR	297
QY	301	ALTAMMAQKVLSPSLPGSILQPVQENSKTSPRSHRSDGEGFPHSHYSDSDEAS	354
Db	298	ALTAMMAQKDLSPSLPGILQPVQEDNSKTSPRSHRSDGEGFPHSHYSDSDEAS	351
RESULT 9			
US-10-004-717-40			
; Sequence 40, Application US/10004717			
; Publication No. US20020192665A1			
; GENERAL INFORMATION:			
; APPLICANT: ZOGHBI, HUDA Y.			
; APPLICANT: YANG, QI			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN			
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,			
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION			
; FILE REFERENCE: P01899US4			
; CURRENT APPLICATION NUMBER: US/10/004,717			
; CURRENT FILING DATE: 2002-08-16			
; PRIOR APPLICATION NUMBER: 09/585,645			
; PRIOR FILING DATE: 2000-06-01			
; PRIOR APPLICATION NUMBER: 60/176,993			
; PRIOR FILING DATE: 2000-01-19			
; PRIOR APPLICATION NUMBER: 60/137,060			
; PRIOR FILING DATE: 1999-06-01			
; NUMBER OF SEQ ID NOS: 69			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 40			
; LENGTH: 266			
; TYPE: PRT			
; ORGANISM: Zebra Fish			
US-10-004-717-40			
Query Match 31.3%; Score 582; DB 13; Length 266;			
Best Local Similarity 46.5%; Pred. No. 8, 1e-37;			
Matches 155; Conservative 21; Mismatches 69; Indels 88; Gaps 14;			
QY	12	EVKELGDHHRQPHHLFPQPPPPPPATLQAREHPVVPPELSLSDSDPPRAWLAPVQAG	71
Db	10	EVVELDVQH-----SSLRGEQSKYPPALMALMASSDPPRAWLAPVQAG	51
QY	72	ICTARAAQYLLHSPELGASEAAAPRDEVDGREGELVRRSSGASSKSPGVKVBQLCKL	131
Db	52	TCAAH-AEYLLHSP--GSS-----AEGVSSASNFKSKS--PVKVR-LCRL	93
QY	132	KGVVVDLGCQRAPSSKQVNGVKORLAAANARERRRMHGLNHAFLDRLNVPSPNN	191
Db	94	KGAVGADE---GRQAPSSKSTNVVQKORMAANARERRRMHGLNHAFLDRLSVIPADN	150
QY	192	DKKLSKYETLQMAQIYINALSELLQTPSGGEGPPPPPPASCKSDHHH-LRTAASYEGGAGN	250
Db	151	DKKLSKYETLQMAQIYINALSDLLQGPCAKADPP---NCDLLHANVLELTDSPRGS---	203
QY	251	ATAAGAAQAGSGQRTPPGSCRTFRFSAPASAGGYSVQDLALHFSFTFDSALTAMMAQXN	310
Db	204	-----PGVCR-----GTGVGYPYQ-----YEDGTFSFMEQDL	232
QY	311	LSP---SLFSGILQPVQENSKTSPRSHRSDGE	340
Db	233	QSPSGTSKSGS-----EASKDSPRNSRSDGE	258
RESULT 10			
US-10-004-717-66			
; Sequence 66, Application US/10004717			
; Publication No. US20020192665A1			
; GENERAL INFORMATION:			
; APPLICANT: ZOGHBI, HUDA Y.			
; APPLICANT: YANG, QI			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN			

```
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P018990U84
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Frog
US-10-004-717-66

Query Match      28.8%; Score 537; DB 13; Length 259;
Best Local Similarity 41.8%; Pred. No. 2.3e-33; Indels 128; Gaps 16;
Matches 154; Conservative 19; Mismatches 67;

QY 1 MSRLH-----ADEMAEVKELGDHHRQPOPHLPQPPPPPPATLQAREHPVYPPELSLL 56
Db 1 MARLHGATAADWCXLKEL-----PSE--AGLLARDY-----LL 33

QY 57 DSTDPRLMPT-LQIGICTARAQYLLHSPGLGASEAAAPRDVDRGRLVRRSSGASS 115
Db 34 DSSDPRLMPT-SLQS-----RPEVYLHPP-----GRAHKVR----- 65

QY 116 SKSPGPKVKREOLCKLKGVVVDE-----LGCSPQAPSSKQVNGVOKQRL 162
Db 66 -----SLCKLKLGRDDDDDEDEDEERSEGLCHRRGPPGKPGVQKQRL 114

QY 163 ANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSELLQTPSGGE 222
Db 115 ANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSELLQTPSGGE 174

QY 223 QPPPPASCKSDHHLRTAAVEGGAGNATAAGAAQASGGSGRPTPPGSCRTFRFAPASA 282
Db 175 DPECPPT-----YQLHSGPEPRLVQSGSC-MRFS----- 202

QY 283 GYSVOLDALHSTFEDSDALTAMAKNLSPSLPGSILQPVQENSKTSPRHSRSGERS 342
Db 203 GDFPGQ-SPLSFQFGAALSG---KGI-----GSAPSSSGEDSKTSPRHSRSGEFT- 251

QY 343 PHSYSDS 350
Db 252 -RSPYSES 258

RESULT 11
US-10-004-717-60
; Sequence 60, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS.
; TITLE OF INVENTION: OSTEOARTHRITIS, AND ABNORMAL CELL PROLIFERATION
; FILE REFERENCE: P018990U84
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 60
; LENGTH: 161
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-60

Query Match      18.1%; Score 337; DB 13; Length 161;
Best Local Similarity 61.2%; Pred. No. 3.6e-18; Indels 18; Gaps 5;
Matches 79; Conservative 10; Mismatches 22;

QY 107 RRSGGAS---SSKSPGPKVKREOLCKLKGVVVDELGCSQAPSSKQVNGVOKQRLA 163
Db 42 RPVAGGARVPAGAPGE---RGAAGARGGG-----GGAGPRA---QVSGVQKQRLA 89

QY 164 ANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQ 223
Db 90 ANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSELLQTPSGGEQ 146

QY 224 PPPPPASCK 232
Db 147 PPEPPAKAE 155

RESULT 12
US-08-722-570-8
; Sequence 8, Application US/08722570
; Publication No. US20030044887A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/722,570
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 5365
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-722-570-8

Query Match      15.1%; Score 281; DB 8; Length 57;
Best Local Similarity 96.5%; Pred. No. 2e-14; Indels 0; Gaps 0;
Matches 55; Conservative 1; Mismatches 1;

QY 158 KQRLAANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSELL 214
Db 1 KNRLAANARERRRMHGLNHAFOQLRNVPISFNNDKLSKYETLQMAQIYINALSEI 57
```

```
RESULT 13
US-10-425-259-8
; Sequence 8, Application US/10425259
; Publication No. US20030224431A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
;           Ma, Qiufu
; TITLE OF INVENTION: NEUROGENIN
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/425,259
; FILING DATE: 29-Apr-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,411A
; FILING DATE: 15-SEP-1997
; APPLICATION NUMBER: US 08/772,009
; FILING DATE: 19-DEC-1996
; APPLICATION NUMBER: US 08/722,570
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-63902-3/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-425-259-8
Query Match 15.1%; Score 281; DB 12; Length 57;
Best Local Similarity 96.5%; Pred. No. 2e-14; Indels 0; Gaps 0;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 158 KORLAANARERRMEGLNHAFLQRLNVPSPFNNDKKLSKYETLQMAQIYNALSEL 214
Db 1 KNRRLAANARERRMEGLNHAFLQRLNVPSPFNNDKKLSKYETLQMAQIYNALSEL 57

RESULT 14
US-10-004-717-19
; Sequence 19, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
```

```
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 178
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-19
Query Match 12.3%; Score 228.5; DB 13; Length 178;
Best Local Similarity 34.7%; Pred. No. 9.7e-10;
Matches 77; Conservative 26; Mismatches 54; Indels 65; Gaps 11;

QY 109 SSGGASSKSPGPVKVREQLCKLKGVVVDLGCGRQAPSSKQVGVQKORRLAANARE 168
Db 7 SSGGV--SEPPGAPRERR---RRRG-----RARARTEALLHTLKSRRYKANDRE 51

QY 169 RRRMHGLNHAFLQRLNVPSPFNNDKKLSKYETLQMAQIYNALSELQTPSGGSEQ--PPP 226
Db 52 RNRMHHLNAALDELRSVLPTFPDDTKLTKETLRFAYNIWALSETLRL---AQQLPPP 108

QY 227 P-----PASCCKDHHHLRTAASYEGGAGNATAAGAAQQAQSGGSGRTPPGSCRTRESAP 279
Db 109 PAFRGPPAPSPGSD-----AGSWLSSGS---PAAPSLC----- 139

QY 280 ASAGGYSVQLDALHFTFEDSALT---AMMAQKNLSPSLPGS 318
Db 140 ASAGSPSP-----ATSEDCGVPSDALRAFRGLPAAFGA 175

RESULT 15
US-10-004-717-21
; Sequence 21, Application US/10004717
; Publication No. US20020192665A1
; GENERAL INFORMATION:
; APPLICANT: ZOGHBI, HUDA Y.
; APPLICANT: YANG, QI
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC USE OF AN
; TITLE OF INVENTION: ATONAL ASSOCIATED SEQUENCE FOR DEAFNESS,
; FILE REFERENCE: P01899US4
; CURRENT APPLICATION NUMBER: US/10/004,717
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/585,645
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/176,993
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/137,060
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 213
; TYPE: PRT
; ORGANISM: chicken
US-10-004-717-21
Query Match 12.1%; Score 224.5; DB 13; Length 213;
Best Local Similarity 29.7%; Pred. No. 2.5e-09;
Matches 78; Conservative 24; Mismatches 88; Indels 73; Gaps 9;

QY 32 PPPQPAPATQAREHPVYPPELSLDSTDPRAWLAFTLQICITARAQYLLHSPGLGASE 91
Db 2 PVKAESPA-----PAAEDELLRLASP---AP-----SASLPSSA 34

QY 92 AAAPRDEVDRGELVRRSSGGSSKSPGPVKVREQLCKLKGVVVDLGCGRQAPSSK 151
Db 35 GDEDEDEDGRPR--RLOEGARRAGRGQRPFA-----ARTAE 70
```


Qy	152	QVNGVOKORRLAANAHERRMHGLNHFDPQLRNVI	PSFNDKLSKYETLOMAQIYINAL	211
Db	71	TAQRIKRSRRLKANNRRNRNMNLAALDALRDVLP	TFPEDAKLTKTETLRFHNYIWA	130
Qy	212	SELLQTPS----	GGEQPPPPASCKSDHHHLRTAASYEGGAGNATAAG-----A	256
Db	131	TETLRLAGARLGGAADAFGAAAGSP----	SPASSWSGGASPASPACTLSEGS	187
Qy	257	QAASGGSQRP	PPGSCRTFSAP	279
Db	188	GSASDAEHWP	PPRG----RFAPP	206

Search completed: September 21, 2004, 21:40:26
Job time : 163.424 secs

Blank
Sheet